

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic searchy using sw model

Run on: January 31, 2005, 10:52:07 ; Search time 85 Seconds
(without alignments)
9482.760 Million cell updates/sec

Title: US-10-048-071-27
Perfect score: 1134
Sequence: 1 atgattcaatttcaattaa.....ttaccaccagtcagcaaac 1134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	849	74.9	3200	3	US-09-381-862-2
2	608.2	53.6	1137	4	US-09-583-110-1744
3	608.2	53.6	21338	4	US-08-961-527-20
4	317	28.0	1155	4	US-09-107-532A-902
5	247.4	21.8	1158	3	US-09-134-001C-2329
6	236.2	20.8	2347	4	US-08-956-171E-153
7	236.2	20.8	2347	4	US-08-781-986A-153
8	159.6	14.1	546	4	US-09-134-000C-1634
9	59.2	5.2	640681	4	US-09-790-988-1
10	57.2	5.0	1141	4	US-09-808-708B-22
11	54.8	4.8	1141	4	US-09-806-708B-22
12	49.6	4.4	426	4	US-09-134-000C-1635
13	49	4.3	640681	4	US-09-790-988-1
14	47.4	4.2	14066	4	US-09-601-198-56
15	46.4	4.1	471	4	US-09-543-681A-1232
16	45.8	4.0	751	4	US-08-956-171E-892
17	45.8	4.0	751	4	US-08-781-986A-892
18	45.6	4.0	832	4	US-09-621-976-2813
19	45.4	4.0	821	3	US-08-998-416-541
20	45	4.0	1134	4	US-09-601-198-62
21	44.8	4.0	423	4	US-09-710-279-1771
22	44.8	4.0	861	4	US-09-710-279-1145
23	44.8	4.0	984	3	US-09-134-001C-2705
24	44.8	4.0	3315	4	US-09-710-279-3820
25	44.8	4.0	3801	4	US-09-710-279-4271
26	44.8	4.0	3926	4	US-09-710-279-4300
27	43.8	3.9	744	4	US-09-248-796A-778

28	43.4	3.8	15598	4	US-08-956-171E-82	Sequence 82, Appl
29	43.4	3.8	15598	4	US-08-781-986A-82	Sequence 82, Appl
30	43.4	3.8	1664976	4	US-08-916-421B-1	Sequence 1, Appl
31	43.4	3.8	1664976	4	US-09-692-570-1	Sequence 13, Appl
c	42.8	3.8	6113	4	US-10-204-708-13	Sequence 208, App
c	42.4	3.7	3001	4	US-09-539-333D-208	Sequence 6131, Ap
33	42.2	3.7	2919	4	US-08-248-796A-6131	Sequence 5237, Ap
34	41.6	3.7	1782	4	US-09-893-600-1	Sequence 11532, A
c	41.6	3.7	2517	4	US-09-248-796A-11532	Sequence 44, App
37	41.4	3.7	603	4	US-09-248-796A-444	Sequence 167, App
38	41.4	3.7	825	4	US-09-806-708B-23	Sequence 3017, Ap
c	41.4	3.6	1055	4	US-09-601-198-167	Sequence 3, Appl
40	41.2	3.6	741	4	US-07-715-184-3	Sequence 27, Appl
c	41.2	3.6	1347	4	US-07-876-280-6	Sequence 1, Appl
42	41	3.6	1425	1	US-07-876-280-6	Sequence 1, Appl
43	41	3.6	1425	1	US-07-876-280-6	Sequence 1, Appl
44	41	3.6	1425	1	US-07-876-280-6	Sequence 1, Appl
45	41	3.6	1425	1	US-07-935-310A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-381-862-2
; Sequence 2, Application US/09381862
; Patent No. 6245906
; GENERAL INFORMATION:
; APPLICANT: Ueyama, Hiroshi
; APPLICANT: Abe, Kanako
; APPLICANT: Keshi, Hiroyuki
; APPLICANT: Matsuhisa, Akio
; TITLE OF INVENTION: PROBES FOR THE DIAGNOSIS OF INFECTIONS
; CAUSED BY STREPTOCOCCUS PYOGENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/381,862
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: JP 1997-71077
; FILING DATE: 25-MAR-1997
; APPLICATION NUMBER: PCT/JP98/01289
; FILING DATE: 23-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Cawley, Jr., Thomas A.
; REGISTRATION NUMBER: 40,944
; REFERENCE/DOCKET NUMBER: 19036/36274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes

STRAIN: Clinical Isolate SP-7-44
US-09-361-862-2

Query Match 74.9%; Score 849; DB 3; Length 3200;
Best Local Similarity 98.7%; Pred. No. 9.5e-204;
Matches 866; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 1 ATGATTCAATTTTCAATTAATGCGACATATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 2324 ATGATTCAATTTTCAATTAATGCGACATATTTATTCATGCTTTTAAATGCAACTAAACGT 2383
QY 61 GCTATTACACTAAATGCGCATTCCTATTCTTTCATCAATAAATAAATGGAAGTCACTTCT 120
DB 2384 GCTATTACACTAAATGCGCATTCCTATTCTTTCATCAATAAATGGAAGTCACTTCT 2443
QY 121 ACAGAGTAACCTTTAAACAGGCTCTAACCGGTCAATATCAATATGAAACACATATTCCTGTA 180
DB 2444 ACAGAGTAACCTTTAAACAGGCTCTAACCGGTCAATATCAATATGAAACACATATTCCTGTA 2503
QY 181 AGTAATGAATGCTGGTTTGTCTAATTAACCTCTCCAGGAGCTATTTTATTAGAGCTAGT 240
DB 2504 AGTAATGAATGCTGGTTTGTCTAATTAACCTCTCCAGGAGCTATTTTATTAGAGCTAGT 2563
QY 241 TTTTATTAATTAATTAATTTCAAGTTTGGCAGATATTAGTATAAATGTTTAAAGAAATTGAA 300
DB 2564 TTTTATTAATTAATTAATTTCAAGTTTGGCAGATATTAGTATAAATGTTTAAAGAAATTGAA 2623
QY 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAGATGTT 360
DB 2624 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAGATGTT 2683
QY 361 GACAGTATCCTGCTACAGAGATATCAACAGAAATCCTTTGATTTTAAACCAAAA 420
DB 2684 GACAGTATCCTGCTACAGAGATATCAACAGAAATCCTTTGATTTTAAACCAAAA 2743
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTGCCT 480
DB 2744 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTGCCT 2803
QY 481 ATTTTAAACAGGAGTTCATATTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
DB 2804 ATTTTAAACAGGAGTTCATATTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 2863
QY 541 GACTCTCATGCTATGAGCCAAAGTTTAAATCACTTTGGAC-AAATCTTCAGCAGATTTCAT 599
DB 2864 GACTCTCATGCTATGAGCCAAAGTTTAAATCACTTTGGACAAATCTTCAGCAGATTTCAT 2923
QY 600 GGTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATTTGA 659
DB 2924 GGTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATTTGA 2983
QY 660 GACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATTTCTTT 719
DB 2984 GACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATTTCTTT 3043
QY 720 TTATACACGCTCTTAGAAGGAAATATCCCGATACAGACCGCTTTTAAATGACAGAAAT 779
DB 3044 TTATACACGCTCTTAGAAGGAAATATCCCGATACAGATCGTTTATTAATGACAGAAAT 3103
QY 780 TGAGACGAGGTGTTTTCATACCCAAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 839
DB 3104 TGAGACGAGGTGTTTTCATACCCAAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 3163
QY 840 GATTCTTAATGCTACTCAAAATGGTACTGTTAAGCTT 876
DB 3164 GATTCTTAATGCTACTCAAAATGGTACTGTTAAGCTT 3200
```

RESULT 2

US-09-583-110-1744
; Sequence 1744; Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 1744
LENGTH: 1137
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-583-110-1744

Query Match 53.6%; Score 608.2; DB 4; Length 1137;
Best Local Similarity 71.1%; Pred. No. 2e-143;
Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

```
QY 1 ATGATTCAATTTTCAATTAATGCGACATATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 1 ATGATTCAATTTTCAATTAATGCGACATATTTATTCATGCTTTTAAATACAACTAAACGT 60
QY 61 GCTATTAGCACTAAATGCGCATTCCTATTCTTTCATCAATAAATAAATGGAAGTCACTTCT 120
DB 61 GCTATTAGCTTAAATGCGCATTCCTATTCTTTCATCAATAAATAAATGGAAGTCACTTCT 120
QY 121 ACAGAGTAACCTTTAAACAGGCTCTAACCGGTCAATATCAATATGAAACACATATTCCTGTA 180
DB 121 GAAGTATTAATTTAAATGCGTCAATATGGAATTTCAATTTGAAATTTTATTTCTCAA 180
QY 181 AGTAATGAATGCTGGTTTGTCTAATTAACCTCTCCAGAGCTATTTTATTAGAGCTAGT 240
DB 181 AAAAATGAAGTGGTGGTTGTTAAATTAATTTCTTTAGGTTGATCTCTTTGAAAGCTTCT 240
QY 241 TTTTATTAATTAATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATTGAA 300
DB 241 TTTTATTAATTAATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATTGAA 300
QY 301 CAACACCAAGTGTGTTTAAACAGTGGTAAATCAGAGATTAACCTTTAAAGGAAAGATGTT 360
DB 301 CAAATCAAATGTTTAAACAGTGGTAAATCAGAGATTAACCTTTAAAGGAAAGATGTT 360
QY 361 GACAGTATCCTGCTCTACAGAGATATCAACAGAAATCCTTTGATTTTAAACCAAAA 420
DB 361 GAACAATATCCAGAAATCCAGAAATTTTCAGCAAGCACTCTCTTTAACTTTGAAACAAA 420
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTGCCTCT 480
DB 421 TTACTCAGAAATTAATTAATGAACAGCTTTTGTCTGCAAGTACACAGAGAGTGCCTCG 480
QY 481 ATTTTAAACAGAGTTCATATTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
DB 481 ATTTTAAACAGAGTTCATATTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
QY 541 GACTCTCATGCTATGAGCCAAAGTTTAAATCACTTTGGCAATATCTTCAGCAGATTTCATG 600
DB 541 GACTCTCATGCTATGAGCCAAAGTTTAAATCACTTTGGCAATATCTTCAGCAGATTTCATG 600
QY 601 GTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTCAG 660
DB 601 GTCTTAATTCCTAGCGGTTCTCTACGCGAATTTTTCAGCGGATTTTACAGATGATATTCAG 660
QY 661 ACCGTTGAGGTATTTTCTCCAGGCGAATCTTGTTCAGAGTGAACACATTTCTTTT 720
DB 661 ACTGTAGAGATTTTCTTTGCGCAATTAACCAAAATCTCTTTTAGAGCGGAAATATTAGCTTC 720
QY 721 TATACACGCTCTTTAGAGAGGAAATTTATCCCGATACAGACCGCTTTTATTAATGACAGAAAT 780
DB 721 TATACACGCTCTTTAGAGAGGAAATTTATCCCGATACAGACCGCTTTTATTAATGACAGAAAT 780
```

Db 721 TATACTCGTCTCTAGAGGAACTATCTCTGATACAGATCGCTTGAATCCAAACAGACTTT 780
 Qy 781 GAGACGAGGTTGTTTCAATACCCAAATCCCTTCGCCAGCTATGGAAGCTGCTTCTTG 840
 Db 781 AACACTACTATTACTTTTAAATGGTGAATCTTACGCCAGTCAATGGAGCGTCCCGTCTT 840
 Qy 841 ATTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGTTAGTACTCAAAATCATATTCA 900
 Db 841 TTATCAAGTGGGACTCAAAATGGTACTGTTAAGCTTGAAATTTAAAGATGGGTTGTTAGC 900
 Qy 901 GCTCATGTTAACTCACTCAGCTGAGTGGTAAAGTGAACGAGGATTTAGATATTGTTAGTCAG 960
 Db 901 GCCCATGTTCACTCTCCAGAAGTGGTAAAGTGAACGAGGATTTAGATATTGTTAGTCAG 960
 Qy 961 TCTGTAGTGAATTTAACTATCAGCTTCAATCAAACTTACCTTATTGAGCTTTTAAAGCT 1020
 Db 961 ACTGGTGAAGATTTGACCAATAGTTTCAACCCAACTTACTGATTGTTCTCTAAAGCT 1020
 Qy 1021 ATTTAAAGTGAACGAGTAAATTTTCAATTTTATCAACCAAGTTCAGCTTACTTCTTGTG 1080
 Db 1021 TTAAATAGGAAAGGTGACCAATAGCTTTATCTCAGCTGTTTCTGCTCCATTTTACTCTGTG 1080
 Qy 1081 CCAGCGATGAGGAAAGTTTTATCCAAATTTTATCAACCAAGTTCAGCTTACTTCTTGTG 1133
 Db 1081 CCAGCGATGAGGAAAGTTTTATCCAAATTTTATCAACCAAGTTCAGCTTACTTCTTGTG 1133

RESULT 3

US-08-961-527-20
 ; Sequence 20, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 21338 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-961-527-20

Query Match 53.6%; Score 608.2; DB 4; Length 21338;
 Best Local Similarity 71.1%; Pred. No. 5,7e-143;
 Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

Qy 1 ATGATTCAATTTTCAATTAATCGCACATTAATTTATTCATGCTTTTAAATACAACTAAACGT 60
 Db 7076 ATGATTCAATTTTCAATTAATTAATTAATTTTCTTCAAGCATTAATTAATTAAGAGA 7135
 Qy 61 GCTATTAGCAGTAAATGCGCATTCCTATTCTTTTCATCAATAAATTTGAAGTCACTTCT 120
 Db 7136 GCTATTAGTCTTAAATGCGCATTCCTATTATTAACAGTAAATTTGACGTGACCAAT 7195
 Qy 121 ACAGAGTAACCTTTTAAACAGGCTCTAACGGTCAAAATATCAATTTGAAAAACAATTTCTGTA 180
 Db 7196 GAAGGTATTACTTTTAAATGCTTCAATGCTCAATTTCAATTTGAAAAATTTTATTTCTCA 7255
 Qy 181 AGTAATGAAATGCTGGTTTGTCTAATTAATCTCTCCAGGAGCTATTTTATTAAGAGTAGT 240
 Db 7256 AAAATGAGATGCTGGTTTGTCTAATTAATTTCTTTAGGTTCGATCCTCTTGAAGCTTCT 7315
 Qy 241 TTTTATTAATATTATTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATTTGAA 300
 Db 7316 TTTTATTAATATTATTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATTTGAA 7375
 Qy 301 CAACACCAAGTTTGTTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAAAGATGTT 360
 Db 7376 CAATATCAATTTGTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAAAGATGTT 7435
 Qy 361 GACGATATCTGCTCTAAGAGATATCAACAGAAATCCTTTGATTTTAAAGAAACAAA 420
 Db 7436 GAACAATATCCACGAATCCAGAAATTTTCAAGCAAGCTCTCTTTAATTAATTTGATTTGAT 7495
 Qy 421 TTATTTGAAGTCTATTATTTGCTGAAACAGCTTTTGCAGCAGCTTTTCAAGAAAGTCTGCT 480
 Db 7496 TTACTCAAGAAATTTATTAATGAAACAGCTTTTGCAGCAGCTTTTCAAGAAAGTCTGCTG 7555
 Qy 481 ATTTTAAACAGGATTTTATTAAGTAAATCAATAAAGATTTTAAAGCACTAGCAGCT 540
 Db 7556 ATTTTAAACAGGATTTTATTAAGTAAATCAATAAAGATTTTAAAGCACTAGCAGCT 7615
 Qy 541 GACTCTCATGATGAGCAGCAAGTTTAAATCACTTTTGGACAAATCTTCAAGCAAGATTTGATG 600
 Db 7616 GACTCTCATGATGAGCAGCAAGTTTAAATCACTTTTGGACAAATCTTCAAGCAAGATTTGATG 7675
 Qy 601 GTAGTTCTTCAAGTAAATCTTTGAGAGAAATTTTCAAGCAAGTATTTTCAAGCAAGATTTGAG 660
 Db 7676 GTAGTTCTTCAAGTAAATCTTTGAGAGAAATTTTCAAGCAAGTATTTTCAAGCAAGATTTGAG 7735
 Qy 661 ACCGTTGAGGATTTTCTTCAAGCAAGCAAAATCTTGTTCAGAAAGTGAACACATTTCTTTT 720
 Db 7736 ACTGTAGAGATTTTCTTTCGCAATTAACCAATCTCTTTAGAGCGGAAATATTAGCTTC 7795
 Qy 721 TATACACGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
 Db 7796 TATACACGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 7855
 Qy 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCAGCTATCGAAGCTGCTCTTCTTG 840
 Db 7856 AACACTACTATTACTTTTAAATGTTGTAACCTTTACCCAGTCAATGGAGCGTCCCGCTTT 7915
 Qy 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTTCAAAATCATATTTCA 900
 Db 7916 TTAATCAAGTGGAGTCAAAATGGTACTGTTGAAATTTGAAATTAAGATGGGTTGTTAGC 7975
 Qy 901 GCTCATGTTAACTCACTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCAG 960
 Db 7976 GCTCATGTTAACTCACTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCAG 8035
 Qy 961 TCTGTTAGTGAATTTAACTCATGCTTCAATCCAGTTCCTTATTAAGCTTTTAAAGCT 1020
 Db 8036 ACTGTTAGGATTTGACCAATAGTTTCAACCCAACTTACTTACTTGAATTTCTCTTAAAGCT 8095
 Qy 1021 ATTTAAAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTCGACCAATTTACCCCTAACA 1080
 Db 8096 TTAATACGAAAGGAGTCACTATTAGCTTTTATCTCAGCTGTTTCTGCTCAATTTACTCTTG 8155
 Qy 1081 CCAGCGATGAGGAAAGTTTTATCCAAATTTTATCAACCAAGTTCAGCAACAA 1133

FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 2329
 LENGTH: 1158
 TYPE: DNA
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-2329

Query Match 21.8%; Score 247.4; DB 3; Length 1158;
 Best Local Similarity 52.9%; Pred. No. 8.1e-53;
 Matches 580; Conservative 0; Mismatches 511; Indels 6; Gaps 2;

QY 1 ATGATTCATTTTCAATTAATGCGACATATTTATTTCATGCTTTTAAATACAACTAAACGT 60
 DB 25 ATGATGGAATTCACAAATTAAGAGAGATTATTTTATTAATCACTTTAACGACACATTTAAAG 84
 QY 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTCATCAATAAAAATTTGAAGTCACTTCT 120
 DB 85 GCTATCTCACCTAGAACAACTTTACCGATTTTAACAGGTATCAAAATTTGATGCTTAAGAA 144
 QY 121 ACAGAGTAACCTTTAACAGGGTCTAACGGTCAAAATATCAATTTGAAAAACACTATTC--T 177
 DB 145 AATGAAGTCATTTTAACTGGGTGAGATTCAGAAATATCAATAGAAATCACTATTCCAAAA 204
 QY 178 GTAAGTAATGAATGCTGGTTGCTAATTACCTCTCAGGAGCTATTTTATTAGAAGCT 237
 DB 205 CAAGTTGATGGTGAAGAAATTTGTAATTTACAGAAACAGGATCACTTCTTCTCGT 264
 QY 238 AGTTTTTTTATTAATATTATTTCAAGTTTGCAGATATTAGTATATAATTTTAAAGAAAT 297
 DB 265 CGTTTCTCGTTGATATTATTAATAAACTCTCGAAAGAAAGTTAAATTTCACTAAT 324
 QY 298 GAACAACACCAAGTTGTTTAAACCAAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGAT 357
 DB 325 GAACAATTTCCAAACGTTAATACATCAGTCAGTCACTTCAAGATTTTAACTTAAGTGGCTTAG 384
 QY 358 GTTGACAGTATCTCGTCTACAGAAAGTATCAACAGAAATTCCTTTGATTTTAAACA 417
 DB 385 CCGTATCAATACCCATTTATCTCAGGATATCAGAGATGACGCTATTCAATTTGTCAGTT 444
 QY 418 AAATTTTGAAGTCTATTATTGCTCAACAGCTTTTGCAGCCAGTTTCAACAGAAAGTCT 477
 DB 445 AAGTACTAATAAATATCATGCAAACTAATTTTGCAGTGTCACCTCAGAAACAGCA 504
 QY 478 CCTATTTTAAACAGGAGTTTATTTGATTTAAGTAAATCAATAAGATTTTAAAGCAGTAGCG 537
 DB 505 CCAGTACTTACTGGTAACTGGCTTATA---CAAGATAATGAATTAATATGCACAGCA 561
 QY 538 ACTGACTCTCATCGTATGAGCCAACTTTAATCACTTTTGGCAATTAATTTTACGAGATTG 597
 DB 562 ACAGATTCACACCGCTTAGCTGTAAGAAAGTTTACAGTTTAGAAGATGAATTCAGAAATAA 621
 QY 598 ATGGTGTGTTCTTCCAAAGTAAATCTTTTGAGAGAAATTTTCCAGAGTATTACAGATGAT 657
 DB 622 AATGTCATATCTCGTTGAAGCTTTTATCTGAAATTAACAAATTAATGAGTGACAGGAC 681
 QY 658 GAGACCGTTGAGGATTTTCTTCCAAAGCCAAATCTTTGTTTCAAGAGTGAACACATTTCT 717
 DB 682 GAAGATATTTGATATTTCTTCTTCTTCAACCAAGTGTATTTCAGAGTGGGAATTAAT 741
 QY 718 TTTTATACCGCTCTTGAAGAAATTTATCCCGATACAGCCGTTTATTATATGACAGAA 777
 DB 742 TTTCTCTCACGTTTACTTGAAGGTCAATTTATCCAGATACGACAGCTTTATTTCCAGAAAT 801
 QY 778 TTTGAGCGGAGTTGTTTTCATACCCCAATCCCTTCGCGCAGCTATGGAAGTCGCTTC 837
 DB 802 TATGAGATTAATAGGAATTAACAAATGAGAGACTTCTATCATGCAATTTGATCGTCACT 861

QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTTAAGCTTGAGATTACTCAAAATCATATT 897
 DB 862 TTATTAGCACGTGAAGGTGGAAATAATCTTATTAAATTAAGTACAGGTAATGAATAGTT 921
 QY 898 TCAGCTCATGTAATCACTCACCTGAGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGT 957
 DB 922 GAATCTTTCATCTACTTCTCCTGAAATTTGTTACTGTTTAAAGAAAGAGTTTAAACGCTAAT 981
 QY 958 CAGTCTGGTAGTGATTTAACTATCAGCTTCACTCAACTTCACTTATTTGAGTCTTTTAAAA 1017
 DB 982 GTAGNAGCGGAAACTTTGAAATTTCTTTCAACTCAAAATACATGATGATGCTTTTAAAA 1041
 QY 1018 GCTATTTAAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTTCGACCATTCACCTTA 1077
 DB 1042 GCCATTGATAATGATGAAGTAGAATGATTTCTTTGGTACAAATGAAACCATTTATCTTTA 1101
 QY 1078 ACACNAGCGGATGAGGA 1094
 DB 1102 AAACCAAGATGATGA 1118

RESULT 6

US-08-956-171E-153
 ; Sequence 153, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:

APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5256
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248PI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 153:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2347 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 153:

US-08-956-171E-153

Query Match

20.8%; Score 236.2; DB 4; Length 2347;

Best Local Similarity 52.2%; Pred. No. 6.8e-50;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 393 ATGATGGAAATTCATTTAAAGAGATTTATTTTATACACAAATTAATGACATTTAAA 452

QY 61 GCTATTAGCATTAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB 453 GCTATTTCACCAAGAACAACTTACTATTAATTAACCTGGTATCAAAATCGATCGGAAGAA 512

QY 121 ACAGAGTAACCTTTAAACGGGCTTAACGGTCAATATCAATTTGAAAAACACTATTCCTGTA 180
DB 513 CATGAAGTTATTAATTAAGTCTGAGCTTCAAAATTTCAATGAAATCACTATTCCTAAA 572

QY 181 AGT---AATGAAATGCTGGTTGCTTAATTAACCTCTCCAGGAGCTATTTTATTAGAACT 237
DB 573 ACTGTAGATGGGAAGATATTGTCNAATTTTCAGAAACAGGCTCAGTAGTACTTCCCTGGA 632

QY 238 AGTTTTTTTATTAATTAATTTCAAGTTTTCAGATATTAGTATTAATTTTAAAGAAAT 297
DB 633 CGATTTCTTTGTTGATTTATTAATAAAATTAACCTGGTAAAGATTTAAATTTATCTACAAT 692

QY 298 GAACAACACAGATTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAGAT 357
DB 693 GAACAATTCAGACATTAATTAACATCAGGTCATTTCTGAATTTAATTTAAGTGGCTTAGAT 752

QY 358 GTTGACAGATATCCGCTCAACAAGATATCAACAGAAATCCCTTTGATTTTAAACA 417
DB 753 CCAGATCAATTCCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 812

QY 418 AAATTTTGAAGTCTATTAATTTGCTGAAACAGCTTTTTCAGGCGAGTTTACAAGAAAGTGT 477
DB 813 AAGTGCTTAAACAGCTGATTCACAAACAAATTTTTCAGGCTCCACCTCAGAAACAGC 872

QY 478 CCTATTTTAAACAGGATTCATATTTGTTATTAAGTATCAATAAGATTTTAAAGCAGTAGCG 537
DB 873 CCAGTACTAATCTGGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 929

QY 538 ACTGACTCTCATCTGATGAGCAACGTTTAACTTTTGAACAACTTCTTTCAGGAGATTTG 597
DB 930 ACTGACTCACACCGCTTGGCTGTAAGAAAGTTGAGTGAAGATTTTCTTGAACAA 989

QY 598 ATGGTATGTTTCCAGTAAATCTTTGAGAAATTTTCAGCAGTATTTTACAGATGATAT 657
DB 990 AATGTCATCAATTCAGGTAAGCTTTAGCTGAATTAATAAATTAATGCTGCAATGAA 1049

QY 658 GAGACCGTTGAGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACATTTCT 717
DB 1050 GAAGACATGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1109

QY 718 TTTTATACACGCTCTTGAAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
DB 1110 TTTATTTCTCATATTAAGAGCAATTTCTGATACACAGCTTTATTCCTCCGAAAC 1169

QY 778 TTTGAGCGAGGTTGTTTCAATACCAATCCCTTCCGACGCTATGGAACGTCCTTC 837
DB 1170 TATGAAATTAATTAAGTATAGCAATGGGAGTTTATCATGCTGATTTGCTGCTCT 1229

QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTTCTTAAAGCTTGAGATTTCAAAATCATATT 897
DB 1230 TTATTAGCGGTGAAGTGGTAAATAACGTTTATTAATTAAGTACAGGTGATGACGTTGTT 1289

QY 898 TCAGTCAATGTTAACTCACTGAGGTTGGTAAAGTAAACAGGAGTTTATGATTTGTTAGT 957
DB 1290 GAATTTGCTTCTACATCACCAAGAAATTTGTTACTGTAAGAGAGAGTGTATGCAACGAT 1349

QY 958 CAGTCTGTAGTGAATTAATCAATCAGCTTCAATCAACCTTACCTTATTTAGTCTTTAAA 1017
DB 1350 GTTGAAGGTGTGACCTGAAATTTCAATCACTTAAATATATATGATGATGCTTTAAA 1409

QY 1018 GCTATTAAGAGTGAACAGATTAATTAATTTCTTATCAACAGTTTCAGACATTCACCTTA 1077

DB 1410 GCAATCGATAATGATGAGTTGAAGTTCTTCGGTACATGAAACCATTTATTCTTA 1469

QY 1078 ACACAGCGGATGAGGA 1094
DB 1470 AACCAAAAGGTGACGA 1486

RESULT 7
US-08-781-986A-153
; Sequence 153, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248pp
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-153

Query Match 20.8%; Score 236.2; DB 4; Length 2347;
Best Local Similarity 52.2%; Pred. No. 6.8e-50;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 393 ATGATGAAATTCATTTAAAGAGATTTATTTTATACACAAATTAATGACATTTAAA 452

QY 61 GCTATTAGCATTAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB 453 GCTATTTCACCAAGAACAACTTACTATTAATTAACCTGGTATCAAAATCGATCGGAAGAA 512

QY 121 ACAGAGTAACCTTTAAACGGGCTTAACGGTCAATATCAATTTGAAAAACACTATTCCTGTA 180
DB 513 CATGAAGTTATTAATTAAGTCTGAGCTTCAAAATTTCAATGAAATCACTATTCCTAAA 572

QY 181 AGT---AATGAAATGCTGGTTGCTTAATTAACCTCTCCAGGAGCTATTTTATTAGAACT 237
DB 573 ACTGTAGATGGGAAGATATTGTCNAATTTTCAGAAACAGGCTCAGTAGTACTTCCCTGGA 632

QY 238 AGTTTTTTTATTAATTTTCAAGTTTTCAGATATTAGTATTAATTTTAAAGAAAT 297
DB 633 CGATTTCTTTGTTGATTTATTAATAAAATTAACCTGGTAAAGATTTAAATTTATCTACAAT 692

Query Match
Best Local Similarity 5.2%; Score 59.2; DB 4; Length 640681;
Matches 455; Conservative 0; Mismatches 548; Indels 27; Gaps 4;

QY 4 ATTCAATTTTCAATTAATCGGACATATTTATTAATGCTTTTAAATACAACTAAACGTGCT 63
DB 12549 ATGAAATTTATTCATCAATAATAATTTTAAATTTTAAATTTTACAAAAATTAAGTCGGTTG 12490
QY 64 ATTAGCACTAAATGCAATTCCTTCTTTCATCAATAATAATAATTAAGTACATCTTAC 123
DB 12489 CTGTGAAAAATCTCTCTTCTTCTTATTTAGATATGTTAAATATTAATAAAGCGA 12430
QY 124 GGAGTAATTTAAACAGGCTCTAAACGCTCAATATCAATTTGAAACACATATTCCTGTAAGT 183
DB 12429 ATGTTATCATTTAAACAGGACGAAATTTAGAAATAGATTTAGTTGCTATGATTTCAACTATCA 12370
QY 184 AATGAAATGCTGTTGCTTAATTAACCTCTCCAGAGCTATTTTATTAAGAAGCTAGTTTT 243
DB 12369 ACAGAACACATATCAGGAACCTGCAACGATTTTCAGGCCGAAAACTTTTGA-----T 12319
QY 244 TTTATTAATATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATGAACAA 303
DB 12318 ATTTGCAGAAATCATTAATATTCATCAATATAGAGATGCAATTAACAATTAATAAATG 12259
QY 304 CACCAAGTTGTTTAAACAGTGGTAAATFCAGAGATTAACCTTTAAAGGAAAAAGATGTTGAC 363
DB 12258 CATATTTTCTGGAATAGTCTGTATATATTAACCACTTACCTTATGATAGTTTTC 12199
QY 364 CAGTATCTCTGCTACAGAGATATCAACAGAAATCTTTGATTTTAAACAAAAATTA 423
DB 12198 GTTCATCATGATTTTCATCATATTTTCAGAA-----TTTTTTTATACCTTCAGATAT 12148
QY 424 TTGAAGTCTATTTTGTGCTGAAACAGCTTTTGAGCAGCTTTTACAGAAAGTCTGCTATT 483
DB 12147 TTAATAAATGATGAAAAAATTCATATTTCTATGCTTAAACAGATGTAGTACTAC 12088
QY 484 TTAACAGAGTTCAATTTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACTGAC 543
DB 12087 CTTAATGCTATTTTATTAGAAAAAACAGATCGATCA---CTTTATGCAAGTAGCCACAGAT 12031
QY 544 TCTCATCTGCTAGCCCAAGTTTAACTATCTTTGGACATATCTCAGCAGATTTGATGTA 603
DB 12030 GGATATCTCTAGGAATATCAAAATTTTTTTTAAAGAAAAATATAATTTCTTTTCAATA 11971
QY 604 GTTCTTCCAAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTGAGACC 663
DB 11970 GTTATTCAGAAAGGCGTTATTGAATATATAGATTTTAAATTTTCCAAACACCA 11911
QY 664 GTTGAAGTATTTTCTCACCAGCCAAATCTTTGTTTCAAGTGAACACATTTCTTTTAT 723
DB 11910 ATAAAGTTTGTAGTTGAAAAAATAATATTAGAGTACATATAGAGATCTTATATTACA 11851
QY 724 ACAGCCTCTTAGAGGAAATTTATCCCATACAGACCGTTTATTAATGACAGATTTGAG 783
DB 11850 ACACAAATTAATTGAAGGCAATATCTGATTAAGAGTGTGTTGCTTGAATAAATAAAC 11791
QY 784 ACGGAGTGTGTTTCAATACCCAAATCTCTCGCCACGCTATGGAACGTCCTTCTTGATT 843
DB 11790 AATTTTATCTTTAAATTTCAAACTATTAAAGCAATCATTTGTTACGAGTCTATTTTA 11731
QY 844 TCTAATGCTACTCAAAATGTTACTGTTAAGCTTGAGATTAATCTCAAAATCATATTTACGT 903
DB 11730 TCCCATGAAAAATTTTGTGGAGTAGAAATTCATATCAGAAATGGTCAATTTAAAGTATTA 11671
QY 904 CATGTTAACTCACCTGAGTGTGTTAAGGTAAACAGAGGATTTAGATTTGTTAGTCAGTCT 963
DB 11670 TCTGATATCAAGAAGAG-----AAATAGCAGAAGATAGATTTTAAATTAATTAATTA 11617
QY 964 GGTAGTGAATTTAACTATACGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCTATT 1023
DB 11616 GGAATAACAGTAAAAATATCAATTAATGTTACTACATATAAGAAATATTAATTTCTATT 11557
QY 1024 AAAAGTGAA 1033

Db 11556 ACTAGCGAAA 11547
RESULT 10
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806.708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 5.0%; Score 57.2; DB 4; Length 1141;
Best Local Similarity 10.2%; Pred. No. 4.7e-05;
Matches 110; Conservative 390; Mismatches 571; Indels 5; Gaps 1;
QY 58 CGTGCTATTAGCACTAAAAATGCCATTCCTATCTTTTCATCAATAAAAAATTTGAAGTCACT 117
DB 53 MSKSRKWTWARMYCKYRRWYNNKSRWKGWYKWKYBCANNYSRYHARRWKDKMTAYBM 112
QY 118 TCTACAGAGTAACCTTTAAACAGGCTCTAACCGGTCAATATCAATATGAAACACATTTCTCT 177
DB 113 TMTNKGKGTWRHRYWRWRAABDTVDHHYVTAMNNAWTTTMCMDKDKRTRWKKNNNA 172
QY 178 GTAAGTAATGAAATGCTGGTTTCTTAATACCTCTCCAGGAGCTATTTTATTAGAAGCT 237
DB 173 TGWDDDTKYHWNNGCBVTVMVRYKTDWDNB-----KMNMYGMBWKKWSYDVTYYW 227
QY 238 AGTTTATTTTAAATATATTTTCAAGTTTGCAGATATTTAGTATAAATTTTAAAGAAAT 297
DB 228 WWDMDCKRKVRVWRIRGRMBYVWABTAHRRYNNWBTBAMAYRWTNNNNNNNA 287
QY 298 GAAACACACCAAGTTGTTTAAACAGGCTGTAATATCAGAGTATCTTAAAGGAAAGAT 357
DB 288 MCKRAKYWGNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDKHKKWKSAAAGVYW 347
QY 358 GTTGACAGATATCTCTGCTACAGAGATATCAACAGAAATCTTTGATTTTAAACAA 417
DB 348 NNNNNWYTKARHBAWDMVHSAWKKWHANAHSYRKKWTBYKRTWNNNNNGTTMW 407
QY 418 AAATTTTGAAGTCTATTTATGCTGAAACAGCTTTTTCAGCAGCTTTTCAAGAAAGTCT 477
DB 408 KRWAAWYKMDMDWBGTYNNNNNGRTYVGTWTKWMTYKWKANNCKWRADHKTCTH 467
QY 478 CCTATTTTAAACAGAGTTTCTATTTGTTATTAAGTAAATCAATAAGATTTTAAAGCAGTACG 537
DB 468 NNTTWMKTYNNNCYKSKMTNGSHRBAAAVYTYWMMWRRYAHANNNNNNNNNN 527
QY 538 ACTGACTCTCATCGTATGAGCAACGTTTATCTTCTTGGACAATCTTTCAGCAGATTTG 597
DB 528 KYBVCSKWNWNYAAWTKSSWNTSYRYKWTNNNSWRWSDTRSMGRANNYARABHYGY 587
QY 598 ATGAGTAGTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGAT 657
DB 588 KWNTRWBWSHTWBHBRAGAAHYWBMWYBAKCHMKRAWKYKAKKYAGAGGNNNNNNNN 647
QY 658 GAGACCGTTGAGGTATTTTCTCCACCAAGCCAAATCTTGTTCAGAGTGAACACATTTCT 717

Query Match
Best Local Similarity 44.2%; Pred. No. 0.00015;
Matches 455; Conservative 0; Mismatches 548; Indels 27; Gaps 4;

QY 4 ATTCAATTTTCAATTAATCGGACATATTTATTAATGCTTTTAAATACAACTAAACGTGCT 63
DB 12549 ATGAAATTTATTCATCAATAATAATTTTAAATTTTAAATTTTACAAAAATTAAGTCGGTTG 12490
QY 64 ATTAGCACTAAATGCAATTCCTTCTTTCATCAATAATAATAATTAAGTACATCTTAC 123
DB 12489 CTGTGAAAAATCTCTCTTCTTCTTATTTAGATATGTTAAATATTAATAAAGCGA 12430
QY 124 GGAGTAATTTAAACAGGCTCTAAACGCTCAATATCAATTTGAAACACATATTCCTGTAAGT 183
DB 12429 ATGTTATCATTTAAACAGGACGAAATTTAGAAATAGATTTAGTTGCTATGATTTCAACTATCA 12370
QY 184 AATGAAATGCTGTTGCTTAATTAACCTCTCCAGAGCTATTTTATTAAGAAGCTAGTTTT 243
DB 12369 ACAGAACACATATCAGGAACCTGCAACGATTTTCAGGCCGAAAACTTTTGA-----T 12319
QY 244 TTTATTAATATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATGAACAA 303
DB 12318 ATTTGCAGAAATCATTAATATTCATCAATATAGAGATGCAATTAACAATTAATAAATG 12259
QY 304 CACCAAGTTGTTTAAACAGTGGTAAATFCAGAGATTAACCTTTAAAGGAAAAAGATGTTGAC 363
DB 12258 CATATTTTCTGGAATAGTCTGTATATATTAACCACTTACCTTATGATAGTTTTC 12199
QY 364 CAGTATCTCTGCTACAGAGATATCAACAGAAATCTTTGATTTTAAACAAAAATTA 423
DB 12198 GTTCATCATGATTTTCATCATATTTTCAGAA-----TTTTTTTATACCTTCAGATAT 12148
QY 424 TTGAAGTCTATTTTGTGCTGAAACAGCTTTTGAGCAGCTTTTACAGAAAGTCTGCTATT 483
DB 12147 TTAATAAATGATGAAAAAATTCATATTTCTATGCTTAAACAGATGTAGTACTAC 12088
QY 484 TTAACAGAGTTCAATTTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACTGAC 543
DB 12087 CTTAATGCTATTTTATTAGAAAAAACAGATCGATCA---CTTTATGCAAGTAGCCACAGAT 12031
QY 544 TCTCATCTGCTAGCCCAAGTTTAACTATCTTTGGACATATCTCAGCAGATTTGATGTA 603
DB 12030 GGATATCTCTAGGAATATCAAAATTTTTTTTAAAGAAAAATATAATTTCTTTTCAATA 11971
QY 604 GTTCTTCCAAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTGAGACC 663
DB 11970 GTTATTCAGAAAGGCGTTATTGAATATATAGATTTTAAATTTTCCAAACACCA 11911
QY 664 GTTGAAGTATTTTCTCACCAGCCAAATCTTTGTTTCAAGTGAACACATTTCTTTTAT 723
DB 11910 ATAAAGTTTGTAGTTGAAAAAATAATATTAGAGTACATATAGAGATCTTATATTACA 11851
QY 724 ACAGCCTCTTAGAGGAAATTTATCCCATACAGACCGTTTATTAATGACAGATTTGAG 783
DB 11850 ACACAAATTAATTGAAGGCAATATCTGATTAAGAGTGTGTTGCTTGAATAAATAAAC 11791
QY 784 ACGGAGTGTGTTTCAATACCCAAATCTCTCGCCACGCTATGGAACGTCCTTCTTGATT 843
DB 11790 AATTTTATCTTTAAATTTCAAACTATTAAAGCAATCATTTGTTACGAGTCTATTTTA 11731
QY 844 TCTAATGCTACTCAAAATGTTACTGTTAAGCTTGAGATTAATCTCAAAATCATATTTACGT 903
DB 11730 TCCCATGAAAAATTTTGTGGAGTAGAAATTCATATCAGAAATGGTCAATTTAAAGTATTA 11671
QY 904 CATGTTAACTCACCTGAGTGTGTTAAGGTAAACAGAGGATTTAGATTTGTTAGTCAGTCT 963
DB 11670 TCTGATATCAAGAAGAG-----AAATAGCAGAAGATAGATTTTAAATTAATTAATTA 11617
QY 964 GGTAGTGAATTTAACTATACGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCTATT 1023
DB 11616 GGAATAACAGTAAAAATATCAATTAATGTTACTACATATAAGAAATATTAATTTCTATT 11557
QY 1024 AAAAGTGAA 1033

Db 11556 ACTAGCGAAA 11547
RESULT 10
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806.708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 5.0%; Score 57.2; DB 4; Length 1141;
Best Local Similarity 10.2%; Pred. No. 4.7e-05;
Matches 110; Conservative 390; Mismatches 571; Indels 5; Gaps 1;
QY 58 CGTGCTATTAGCACTAAAAATGCCATTCCTATCTTTTCATCAATAAAAAATTTGAAGTCACT 117
DB 53 MSKSRKWTWARMYCKYRRWYNNKSRWKGWYKWKYBCANNYSRYHARRWKDKMTAYBM 112
QY 118 TCTACAGAGTAACCTTTAAACAGGCTCTAACCGGTCAATATCAATATGAAACACATTTCTCT 177
DB 113 TMTNKGKGTWRHRYWRWRAABDTVDHHYVTAMNNAWTTTMCMDKDKRTRWKKNNNA 172
QY 178 GTAAGTAATGAAATGCTGGTTTCTTAATACCTCTCCAGGAGCTATTTTATTAGAAGCT 237
DB 173 TGWDDDTKYHWNNGCBVTVMVRYKTDWDNB-----KMNMYGMBWKKWSYDVTYYW 227
QY 238 AGTTTATTTTAAATATATTTTCAAGTTTGCAGATATTTAGTATAAATTTTAAAGAAAT 297
DB 228 WWDMDCKRKVRVWRIRGRMBYVWABTAHRRYNNWBTBAMAYRWTNNNNNNNA 287
QY 298 GAAACACACCAAGTTGTTTAAACAGGCTGTAATATCAGAGTATCTTAAAGGAAAGAT 357
DB 288 MCKRAKYWGNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDKHKKWKSAAAGVYW 347
QY 358 GTTGACAGATATCTCTGCTACAGAGATATCAACAGAAATCTTTGATTTTAAACAA 417
DB 348 NNNNNWYTKARHBAWDMVHSAWKKWHANAHSYRKKWTBYKRTWNNNNNGTTMW 407
QY 418 AAATTTTGAAGTCTATTTATGCTGAAACAGCTTTTTCAGCAGCTTTTCAAGAAAGTCT 477
DB 408 KRWAAWYKMDMDWBGTYNNNNNGRTYVGTWTKWMTYKWKANNCKWRADHKTCTH 467
QY 478 CCTATTTTAAACAGAGTTTCTATTTGTTATTAAGTAAATCAATAAGATTTTAAAGCAGTACG 537
DB 468 NNTTWMKTYNNNCYKSKMTNGSHRBAAAVYTYWMMWRRYAHANNNNNNNNNN 527
QY 538 ACTGACTCTCATCGTATGAGCAACGTTTATCTTCTTGGACAATCTTTCAGCAGATTTG 597
DB 528 KYBVCSKWNWNYAAWTKSSWNTSYRYKWTNNNSWRWSDTRSMGRANNYARABHYGY 587
QY 598 ATGAGTAGTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGAT 657
DB 588 KWNTRWBWSHTWBHBRAGAAHYWBMWYBAKCHMKRAWKYKAKKYAGAGGNNNNNNNN 647
QY 658 GAGACCGTTGAGGTATTTTCTCCACCAAGCCAAATCTTGTTCAGAGTGAACACATTTCT 717

Db 648 NNNNNNATCARDYYAASRWYAMAKWYYKYBAANNAYTHANNWGCWNNATDTRRT 707
Qy 718 TTTTATACACGCTCTTGAAGAGAAATATCCGATACAGACCGTTTATTAATGACAGAA 777
Db 708 MWKNNNNNAGTWKNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAK 767
Qy 778 TTTGACAGCGAGGTTGTTTCAATACCAATCCCTCCGACGCTATGGAAGCGCTTC 837
Db 768 TTDKKNNGAYTKYTTTNNNTYGVVNTAARDGHWANNNNNNNNNNNNNNNNNNNNNN 827
Qy 838 TTGATTTCTAATGCTACTCAAAATGCTACTGTTAAGCTTGAGCTTACCAATCATATT 897
Db 828 AYANGTNNNNNNNNNAYAMWNKYYTDDWRBAYTNNNNNNNNNNNNNNNNNNNNNN 887
Qy 898 TCAGCTCATGTTAACTCACCTGAGGTGGTGAAGTAAACGAGATTTAGATATTGTTAGT 957
Db 888 MSDTCDAMKMDATKMNATTYNRGTAWRTNNNNNNNNNNNNNNNNNNNNNNNNNN 947
Qy 958 CAGCTGCTGAGTGAATTAATCACTACGCTTCAATCAACCTTACCTTATGAGTCTTAA 1017
Db 948 AHTWVVCATKTTKGCWNNCTTCRKYKNNCTWYTWMTTTRTTWYAATRWKTNATGSM 1007
Qy 1018 GCTATTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTTCGACCTTCCACCTA 1077
Db 1008 RCNATGKNNNTYTWGKTRWTAIRWATRMKAWKVMATGSMNTNSYARWYKTRAYKG 1067
Qy 1078 ACACGAGCGATGAGAGAAAGTTTATCCAAATTAATACACAGTACGACAA 1133
Db 1068 YNACAWRWGKATCYMTDANAWTACATSMWATHKYNNWHKCKNNNNNNNNNNNN 1123

RESULT 11

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1) .. (1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 4.8%; Score 54.8; DB 4; Length 1141;
Best Local Similarity 12.4%; Pred. No. 0.00019;
Matches 123; Conservative 335; Mismatches 530; Indels 5; Gaps 2;
Qy 130 ACTTTAAGCGGTCTAAGCGTCAATATCAATGAAACACATATTCCTGTAAGTAATGAA 189
Db 1120 ANNNNNNNNGKDWNRMDATKWSATGTAWTTHAKRGATMCWYWTGTNRRCMRYA 1061
Qy 190 AATGCTGTTTCTAATTAACCTCTCCAGGAGCTATTTTATGAGAGCTAGTTTATTT 249
Db 1060 MRTWYTRSNANWSCATKBNWMTKWTATYRTAWYAMWCMAMNNNNNNNNNNNNNNNN 1001
Qy 250 AATATATTTCGAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATGAACACACAA 309
Db 1000 NAMWATTTWAAYAAKWARWAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 941
Qy 310 GTTGTTTAAACGAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGATGTTGACCAT 369

Db 940 NNNNNTTDDRMMWAKNN 881
Qy 370 CCTCGTCTACAGAGTATCAACAGAAATCCTTTGATTTTAAAAACAAAATTTATGAAG 429
Db 880 RTCRTKYNN 821
Qy 430 TCTATTATGCTGAAACAGCTTTTGCACGAGTTTACAAGAAAGTCGCTCTTATTAACA 489
Db 820 SWCNNN 761
Qy 490 GGAGTTCATATTGATTAAGTAATCATATAAGATTTTAAAGCAGTACGACTCTCAT 549
Db 760 WCYATMNTWYDMMTMTTTRNNMTTSTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 701
Qy 550 CGTATGACCCAACTGTTT---AATCACTTTGCAACAATCTTTCAGCAGATTTGATCGTAGT 605
Db 700 NNWGCWNNNTDARRNTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 641
Qy 606 TCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGTATTTACAGATGATATGAGACCGT 665
Db 640 NNSCTCTRMNTMRWTKGDMTVRKVKVRDITCTYVDVWADSWVWYANMMRCRDVY 581
Qy 666 TGAGGTATTTTCTCACCACCACTTCTGTTGAGAGTGAACACATTTCTTTTATATAC 725
Db 580 TRNNTYKSYAHSYWYWSNNAMWYRYSARNWSSMARWTTTRNNNNNNNNNNNNNNNN 521
Qy 726 ACAGCTCTTAGAAGAAATATCCGATACAGACCGTTTATTAATGACAGAAATTTGAGAC 785
Db 520 RHNNNNNTDTRYNNWKKWARBITTVYDSMCNAKSMRGNWRAKMMWAAANNNDAGAMDH 461
Qy 786 GGAGGTGTTTCAATACCCCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTGATTTTC 845
Db 460 WTYWNGNNTMMRRAMKMMMAWCRRAYCCNNNNNNNNNNNNNNNNNNNNNNNNNNNN 401
Qy 846 TAATGCTACTCAAAATGCTGTTTAACTTGAAGTCTGAGATTAATCAATATTTACGCTCA 905
Db 400 NNBAKMYRMVAMMYSRDITNTDMMMTSDWBWHWYTDYTMRAWNNNNNNNNNNNNNN 341
Qy 906 TGTTAACTCACTGAGGTGGTGAAGTAAACGAGGATTTAGATATTG-TTAGTCACTGCTG 964
Db 340 SWMMMDHWNTHCTYGNNTWGSAYBMAAMSMAWAGASBNVYNNWCRWYMGKTNNNN 281
Qy 965 GTAGTGATTTAATCACTACGCTTCAATCACTTACCTTATGAGTCTTTTAAAGCTATTA 1024
Db 280 NKAAYYRTKTVAWNNRYYDTAVWTBKRYKYCYAYBYWYBYWYMGKHHWRRABHR 221
Qy 1025 AAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCTTAACACCA 1084
Db 220 SWNNWVVKCRNKYVSWHYHAMRYBKWABAGCNNNNNNNNNNNNNNNNNNNNNNNN 161
Qy 1085 CGGATGAGGAAGAAAGTTTATCCAAATTAATTA 1117
Db 160 MHHKKGKAAWNTKNTABRDDHBAHVKYTYW 128

RESULT 12

US-09-134-000C-1635
; Sequence 1635, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; CURRENT APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1635
; LENGTH: 426

```
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1635

Query Match      4.4%; Score 49.6; DB 4; Length 426;
Best Local Similarity 62.5%; Pred. No. 0.0027;
Matches 95; Conservative 0; Mismatches 54; Indels 3; Gaps 1;

Qy 421 TTTATGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAAAGAAAGTCTCT 480
Db 259 TTTGGGAGAAAATATTTTGGACCCGGGTGGTTCATGCCCGAAAGTCTGCCC 318

Qy 481 ATTTAACAGAGTTCATATTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCCGACT 540
Db 319 ATTTTAACTGGGTTCACCTTTATTAGAAAATCAAAAATTACTT---GCCGTTGCGACA 375

Qy 541 GACTCTCATCTATGAGCCCAACGTTTAAATCAC 572
Db 376 GATTACATCGTTTAAAGTCAACGTTGTGATCCC 407

RESULT 13
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, NASHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match      4.3%; Score 49; DB 4; Length 640681;
Best Local Similarity 53.4%; Pred. No. 0.054;
Matches 103; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 246 TATTAATATTATTTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATTTGAACRACA 305
Db 566908 TAATATTATTTCTTTGAAAATAAAAAAATCATTTATACTCCAAAACATGAAATGTT 566967

Qy 306 CCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTTGACCA 365
Db 566968 AAATATTATTATTAATCAATATATTTCTATATAGTATATTTTTTAAAGAAATTTTAAATA 567027

Qy 366 GTATCCTCGTCTACAAGAGTATCAACAGAAAATCCTTTGATTTTAAAAACAAAATTTATT 425
Db 567028 TTTTAAAAATACAAATTAATAATTAATTAATACCTTTCTTCAAAAAAAAATTAAT 567087

Qy 426 GAAGTCTATTATT 438
Db 567088 TCATTTTATTATT 567100

RESULT 14
US-09-601-198-56/c
; Sequence 56, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
```

```
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 56
; LENGTH: 14066
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-56

Query Match      4.2%; Score 47.4; DB 4; Length 14066;
Best Local Similarity 47.5%; Pred. No. 0.034;
Matches 141; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

Qy 3 GATTCAATTTTCAATTAATCGCACATTTATTTCATGCTTTTAATACAACTAAACGTCG 62
Db 2575 GATTAAATGATAAGAAAATATAAGTTTAACTGAAACTGGAATCCGGTTTAAAGT 2516

Qy 63 TATTAGCACTAAAAATGCCATTCTCTATTCTTTCATCAATAAAAAATTTGAAGTCATCTTAC 122
Db 2515 TATTCAAACTCAAAATGATCTATTATGATCTCAACAACTATTATGTAAGTTATC 2456

Qy 123 AGGAGTAACCTTTAAACAGGGTCTAAACGGTCAAAATATCAATTTGAAAAACACTATTCTGTAAG 182
Db 2455 AGGAGTTAACTCTAAATATAATGACGTCAGATTAAAGTAGTCTTTATAAGATAATAATA 2396

Qy 183 TAATGAAAATGCTGGTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 242
Db 2395 TGTATCTATGAAAGTTCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2336

Qy 243 TTTTATTAATATTATTTCAGTTTGCAGATATTAGTATATAATGTTAAAGAAATTTGA 299
Db 2335 GTTATCAATTTAAATTCAAATCGAGATATAGTTTGAAGAAATTTGAATATAATCA 2279

RESULT 15
US-09-543-681A-1232
; Sequence 1232, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1232
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1232

Query Match      4.1%; Score 46.4; DB 4; Length 471;
Best Local Similarity 57.6%; Pred. No. 0.018;
Matches 83; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 159 AATTGAAAACACTATTCCTGTAAGTAATGAAAATGCTGGTTTGCCTTAATTAATTAATTAAT 218
Db 3 AATCGGAAGTAATATGAGTATAAATTTAGACGAAAATATTATACCTGTAAGTCTATTGG 62

Qy 219 AGCTATTTTATAGAGCTAGTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 278
```

Db 63 TAGTATTTCAAGTGGATGATGTTAATGATATATTAAATGCATTGAGCGGAATATTG 122
Qy 279 TATAAATGTTAAAGAAATTTGAACA 302
Db 123 TATAAATGTTATAGAATTTGAAAA 146

RESULT 16
US-08-956-171E-892
; Sequence 892, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 892:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 892:
US-08-956-171E-892

Query Match 4.0%; Score 45.8; DB 4; Length 751;
Best Local Similarity 48.6%; Pred. No. 0.03;
Matches 125; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 72 TAAAAATGCCATTCCTATTCTTTTCATCAATAAAAAATTGAAGTCATCTTCAAGGATGTAAC 131
Db 290 TACAAGGCAAGCATATCGTCTCTCAATTTGGATTTCGATTGAAACATAAAAAAGATGC 349

Qy 132 TTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACACTATTCTGTGAAGTAAATGAAA 191
Db 350 TTTAGCATTTAGAAAAAGCGAAAAATAAAGTTGATAAATCTATTGAAACAAGAGTGAAGC 409

Qy 192 TCGTGGTTTGCTAAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTTTTATTAA 251
Db 410 GATAGCTCATATCAAGTTTAAACCGGATATTTATTATTGTAACATCATTTTTTAGGTAT 469

Qy 252 TATTATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACACCAAGT 311
Db 470 TACATTTCTTGATTGCTGTATGTTGCTATATATACATAAAGCAATAGATGAAACCGAAGA 529

Qy 312 TGTTTAAACCAAGTGGTA 328
Db 410 GATAGCTCATATCAAGTTTAAACCGGATATTTATTATTGTAACATCATTTTTTAGGTAT 469

Qy 252 TATTATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACACCAAGT 311
Db 470 TACATTTCTTGATTGCTGTATGTTGCTATATATACATAAAGCAATAGATGAAACCGAAGA 529

Qy 312 TGTTTAAACCAAGTGGTA 328
Db 530 TGAGTTAGAGAAATTATA 546

RESULT 17
US-08-781-986A-892
; Sequence 892, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 892:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-892

Query Match 4.0%; Score 45.8; DB 4; Length 751;
Best Local Similarity 48.6%; Pred. No. 0.03;
Matches 125; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 72 TAAAAATGCCATTCCTATTCTTTTCATCAATAAAAAATTGAAGTCATCTTCAAGGATGTAAC 131
Db 290 TACAAGGCAAGCATATCGTCTCTCAATTTGGATTTCGATTGAAACATAAAAAAGATGC 349

Qy 132 TTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACACTATTCTGTGAAGTAAATGAAA 191
Db 350 TTTAGCATTTAGAAAAAGCGAAAAATAAAGTTGATAAATCTATTGAAACAAGAGTGAAGC 409

Qy 192 TCGTGGTTTGCTAAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTTTTATTAA 251
Db 410 GATAGCTCATATCAAGTTTAAACCGGATATTTATTATTGTAACATCATTTTTTAGGTAT 469

Qy 252 TATTATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACACCAAGT 311
Db 470 TACATTTCTTGATTGCTGTATGTTGCTATATATACATAAAGCAATAGATGAAACCGAAGA 529

Qy 312 TGTTTAAACCAAGTGGTA 328
Db 410 GATAGCTCATATCAAGTTTAAACCGGATATTTATTATTGTAACATCATTTTTTAGGTAT 469

Db 530 TGAGTTAGAGAAATTATA 546

RESULT 18

US-09-621-976-2813

; Sequence 2813, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 2813

; LENGTH: 832

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 235..399

US-09-621-976-2813

Query Match 4.0%; Score 45.6; DB 4; Length 832;

Best Local Similarity 14.0%; Pred. No. 0.035; Mismatches 139; Indels 4; Gaps 2;

Matches 51; Conservative 170; Indels 4; Gaps 2;

Qy 420 ATTATTGAAGTCTATTATTGCTGAACAGCTTTTGCAGCCAGTTTACAGAAAGTCGTC 479

Db 13 AKCWTKKWSYNNYKWKYKTYWRWRKXKXKWKYKTYWYWRWYKWKYKKA 72

Qy 480 TATTTTACAGAGGTCATATTGTTATTAAGTAATCATTAAGATTTTAAGCAGTAGCGAC 539

Db 73 MCRTKTKKKKKGGYNNYKWKYKTYWRWRKXKXKWKYKTYWYWRWYKWKYKKA 129

Qy 540 TGACTCTCATCGTATGAGCCAGGTTTAAATCACTTTTGACAAATCTTCAGCAGATTGAT 599

Db 130 YRKTCTSSKWTWWRKWKATTTWKKTYWATRYWNNWNNWNNWNNWNNWNNWNNW 189

Qy 600 GGTAGTCTTCCAGTAATCTTTGAGAGAAATTTTCAGCAGTATTTCAGATGATATTTGA 659

Db 190 ARKSTWRKRSYASARSAGKCCYSCSWGMSWKYNNWNNWNNWNNWNNWNNWNNW 249

Qy 660 GACGTTGAGTATTTTCTCACCAGCAATCTTGTTCAGAGT-GAACACATTTCTT 718

Db 250 RRYAGSKTSYKSNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNW 309

Qy 719 TTTATACAGGCTCTTAGAAGGAATTTATCCGATACAGACGTTTATTATTAATGACAGAT 778

Db 310 WNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNWNNW 369

Qy 779 TTGA 782

Db 370 ATAA 373

RESULT 19

US-08-998-416-541

; Sequence 541, Application US/08998416

; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippsen, Peter

; APPLICANT: Fohlmann, Rainer

; APPLICANT: Steiner, Sabine

; APPLICANT: Mohr, Christine

; APPLICANT: Wendland, Jürgen

; APPLICANT: Knechtle, Philipp

; APPLICANT: Reibschung, Corinne

; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPH

; AND USES THEREOF

; NUMBER OF SEQUENCES: 1152

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6239264artis Corporation

; STREET: 3054 Cornwalis Road

; CITY: Research Triangle Park

; STATE: No. 6239264th Carolina

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/998,416

; FILING DATE: 24-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: CH 0016/97

; FILING DATE: 31-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy

; REGISTRATION NUMBER: 38,241

; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8587

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 541:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 821 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: PAG1378RP

US-08-998-416-541

Query Match 4.0%; Score 45.4; DB 3; Length 821;

Best Local Similarity 44.3%; Pred. No. 0.039; Mismatches 178; Conservative 0; Indels 0; Gaps 0;

Qy 51 AACTAAACGTCCTATTAGCACTAAATAAGCAATTCCTATTCTTCATCAATAAATGA 110

Db 243 AACATATAATAATAAGCAATTAATAATAATAATAATAATAATAATAATAATA 302

Qy 111 AGTCACCTTACAGGAGTAACTTTAAACAGGCTTAACGGTCAATATCAATGAAACAC 170

Db 303 ATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 362

Qy 171 TATTCCTGTAGTAATGAAATGCTGTTGCTAAATACCTCTCCAGGAGCTATTATT 230

Db 363 AGTTACTGTAGGGAACCTGCGAGTGGGCTTATAATAATCTTTAATAATTCCTTA 422

Qy 231 AGAAGCTAGTTTTTTTATTATAATAATAATAATAATAATAATAATAATAATAATA 290

Db 423 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 482

Qy 291 AGAAATTAACACCAAGTTGTTTAAACAGGTTAAATCAGAGATTACCTTAAAGG 350

Db 483 TATAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 542

Qy 351 AAAAGATGTTGACCAAGTATCTGCTTACAAAGAGTATCAACAGAAATCCTTTGAT 410

Db 543 TATGCATATGTTGGTACACTCTAAATAATAATAATAATAATAATAATAATAATA 602

Qy 411 AAAAACAATAATTAAGTGAAGTCTATTATTGCTGAAACAGCTTT 452

Db 603 ATACCATNAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 644

RESULT 20

US-09-601-198-62


```
; Sequence 62, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-62

Query Match      4.0%; Score 45; DB 4; Length 1134;
Best Local Similarity 44.3%; Pred. No. 0.055;
Matches 183; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 60 TCGTATTAGCACTAAAGATGCCATTCCTTATCTTCATCAATAAATAAGTGAAGTCACTTC 119
Db 639 TGAATTTAAATAGTAATTTTGTGTAGATGTTTAAATATCAACCGTAATATCTTTT 698

Qy 120 TACAGAGTAACCTTTTACAGGCTTACGGTCAATATCAATGAAACACATTTCTCTGT 179
Db 699 TAAAGAAATTAATTTTGAAGCAAGATGATTTAAATTTAACTAATGCGCAACTAAAT 758

Qy 180 AGTAATGAATGCTGTTGCTTAATACCTCTCCAGAGCTATTTTATAGAGTAG 239
Db 759 AACTACTAATAGTAGTAAGTTTCACTACAGCTCCAAACCTGTTGAAATATAAAGTGT 818

Qy 240 TTTTATTAATTAATTTTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATGA 299
Db 819 TTTAATGATACAAATTTGATAATCATCCAGTAGTTTAAATTAATTTGAAGTTAATGA 878

Qy 300 ACAACACCAAGTGTGTTTAAACAGTGTGTAATACAGAGATTACCTTAAAGGAAAGATGT 359
Db 879 TAGTGAGAATAATTTTAAAGAAATGATATTTTAAATAATTAATAAAGTAGTAGACC 938

Qy 360 TCACAGATATCTCTCTACAGAGATATCAACAGAAATCCTTTGATTTTAAACAAA 419
Db 939 AAATGAAGTCAATCTTTTGAAGAAACAGTTGCGATTAGCTAATAATTTTGAAGTATTGA 998

Qy 420 ATTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCAGCTTTTACAGAA 472
Db 999 AATTGAAATACAAAAGAAATGAAATGAAATATGAAATCTTAAGTTAGATCA 1051

RESULT 21
US-09-710-279-1771
; Sequence 1771, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1771
; LENGTH: 423

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1771

Query Match      4.0%; Score 44.8; DB 4; Length 423;
Best Local Similarity 50.5%; Pred. No. 0.043;
Matches 109; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 242 TTTTATTAAATATTATTTCAAGTTTGCAGATATTAGTATAAATCTTAAAGAAATGAAC 301
Db 74 TTTTGAACCTTTATAGATGAAATACTATAGTTACTTTAATTACAGTTGTAAATGTTGAAG 133

Qy 302 AACCAAGTGTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTTG 361
Db 134 AATCGAAACAGATGTTTTTACATGTAAACCAAGGACATAGTTTAAACAAATGAAGAGAAG 193

Qy 362 ACCAGTATCTCTGCTCAAGAAGTATCAACAGAAATCCTTTGATTTTAAACAAAAT 421
Db 194 ACAAAATTATCCAGTATTAATCTGAACCTATTGTTAGAACATAATGTAATTTATGAAGTAAAA 253

Qy 422 TATTGAAGTCTATTATTCGCTGAAACAGCTTTTGCAG 457
Db 254 TTGCACATGGTCTTCCTCGAGAAACAGTGGTTTCAG 289

RESULT 22
US-09-710-279-1145
; Sequence 1145, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1145
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1145

Query Match      4.0%; Score 44.8; DB 4; Length 861;
Best Local Similarity 46.5%; Pred. No. 0.056;
Matches 145; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 4 ATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGTGT 63
Db 151 ATTTACTCTCTTATCAAAACGTATATTTAATGTTGAAGTTGAAATTTTAGTTAGAAAAAG 210

Qy 64 ATTAGCACTAAAGATGCCATTCCTTATCTTCATCAATAAATAATTTGAAGTCACTTCTACA 123
Db 211 ATGAAATTTGAAAAAACAATAATTTATATATGTCGAACAAGATGTTAGCGAAGAAATA 270

Qy 124 GGAGTAATCTTAAACAGGGTCTAAACGGTCAAAATATCAATTTGAAAAACATTTCTCTGAAGT 183
Db 271 CTAAATGATTTAGGAATTTTAAAGGAGGAGTTTTTACTCAGATATTGATCCGATATG 330

Qy 184 AATGAAATGCTGTTGTTGCTAATTTACCTCTCCAGAGCTATTTTATTTAGAACATGTTTT 243
Db 331 ATTTAAGATGATGAAATGAAAGAGTATTATTTAAGAGGGGCTTTCTTAGCAGGTGTTCT 390

Qy 244 TTTTATTAATATTATTTCAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACA 303
```

Db 391 GTAAATATCTGAAACATCTTCATATCATCTTGAATTTTTTCAATATGAAGTCA 450

Qy 304 CACCAAGTGTGT 315
| | | | |

Db 451 TCCGAAGGTCTT 462

```

RESULT 23
US-09-134-001C-2705
; Sequence 2705, Application US/09134001C
; Patent NO. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2705
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2705

```

Query Match 4.0%; Score 44.8; DB 3; Length 984;
Best Local Similarity 46.5%; Pred. No. 0.058;
Matches 14; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy	64	ATTAGCACTAAAAATGCCATTCCCTATTCTTTTCATCAATAAAAAATTGAAGTCACCTTCTACA	123
Db	268	ATGAAATTGAAAAAACAATATTTATATATGTCGACAAAGATGTTAGCGAAAGAAATA	327
Qy	124	CGAGTAACCTTTAAACAGGGTCTAACGGTCAATATCAATTGAAACAACTATTCTCTGAAGT	183
Db	328	CTAATGATTTAGGNAITTTTAAAAAGGGAGTTTTTCTACGATATTGATCCGGATATG	387
Qy	184	AATGAAATGCTGGTTGCTAATTACCTCTCCAGGACTATTTTATTAGAAGCTAGTTTTT	243
Db	388	ATTAAGATGATGAAATGAAAAGAAATTAATTAAAGAGGGGCTTTCTTAGCAGTGTTCTCT	447
Qy	244	TTTATTAATATTTATTTCAAGTTTCCAGATATTAGTATAAATCTTTAAAGAAATTTGAACA	303
Db	448	GTAATAATCTGAAACATCTTCATATCATCTTGAAATTTTTTTCACAATATGAAGATCAT	507
Qy	304	CACCAAGTTGTT	315
Db	508	TCCGAAGCTTT	519

```

RESULT 24
US-09-710-279-3820
; Sequence 3820, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3820

```

```

; LENGTH: 3315
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3820

```

Query Match	4.0%;	Score 44.8;	DB 4;	Length 3315;
Best Local Similarity	46.5%;	Pred. No. 0.091;		
Matches 145;	Conservative 0;	Mismatches 167;	Indels 0;	Gaps 0;
QY	4	ATTCAATTTTCAATTAATCGCACATATTTATTCATGCTTTAAATACAACTAACGTCCT	63	
Db	2611	ATTTACTCTCTTATCAACAGTATATTTAATGTTGAAAGTGAATTTTAGTTAGAAAAAG	2670	
QY	64	ATTAGCACTAAAAATGCCATTCCTATTCTTTTCATCAATAAAAATGGAAGTCATCTTCTACA	123	
Db	2671	ATGAAATTGAAAANAACAATATTTATATATGCGAACAAAGATGTTAGCGAAGAATA	2730	
QY	124	GGAGTAACTTTACAGGGTCAACGGTCAAAATATCAATTGAAAACTATTCCTGTAAAGT	183	
Db	2731	CTAAATGATTTAGGAATTTTAAAAAGGGAGTTTTTACTACGATATTGATCCGATATG	2790	
QY	184	AATGAAAAGTCTGGTTTGCTAAATTACCTCTCCAGGAGCTATTTTATTAGNAGCTAGTTT	243	
Db	2791	ATTTAAGATGATGAAATGAAAAGAGTTATTTTAAGAGGGGCTTTCTTTAGCAGTGGTTCT	2850	
QY	244	TTTATTAAATATTATTTTCAAGTTTGCCAGATATTAGTATAAATGTTTAAAGAAATTTGAACAA	303	
Db	2851	GTAAATAAATCCCTGAACATCTTCATATCATCTTGAATTTTTTTCACAAATATGAAGATCAT	2910	
QY	304	CACCAAGTTGTT	315	
Db	2911	TCCGAAGGTCCT	2922	

RESULT 25
US-09-710-279-4271/c
; Sequence 4271, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4271
; LENGTH: 3801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4271

	Query Match	4.0%	Score 44.8;	DB 4;	Length 3801;
	Best Local Similarity	50.5%;	Pred. No. 0.095;		
	Matches 109;	Conservative 0;	Mismatches 107;	Indels 0;	Gaps 0;
QY	242	TTTTTATTAAATATATTCAAGTTCCGAGATTAGTATATAATGTTAAAGAATTTGAAC	301		
Db	3744	TTTTGAACTTTATAGATGAAATACATACTATAGTTACTTTAAATACAGTTGTTAAATGTTGAAG	3685		
QY	302	AACACCAAGTGTGTTTAAACGAGTGGTAAATCGAGATTACCTTAAAGGGAAGAAGATGTTG	361		
Db	3684	AATCGAAAACAGATGTTTACATGGTAAACAGGACATAGTTTTAAACAAATGAAAGAGAAG	3625		
QY	362	ACCAGTATCCTCGTCTACAGAAGTATCAACGAAAATCCTTTGATTTTAAAAACAAAAT	421		

Db 3624 ACAAAATTATCCAGTAACTGAACATAATTTGTAGAACAATAATGTAATATGAAGTAAAAA 3565
Qy 422 TATTGAAGTCTATTATTGCTGAAACAGCTTTTTCGAG 457
Db 3564 TTGCACATGCTCTTCTGCGAAGACAGTGTTTCAG 3529

RESULT 26

US-09-710-279-4300/c
; Sequence 4300, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4300
; LENGTH: 3926
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4300

Query Match 4.0%; Score 44.8; DB 4; Length 3926;
Best Local Similarity 50.5%; Pred. No. 0.096;
Matches 109; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
Qy 242 TTTTATTAATATTTTCAAGTTTCCAGATATTAGTATAATGTTAAAGAAATTGAC 301
Db 985 TTTTGAACCTTTATAGATGAATACTATAGTTACTTTAATACAGTTGTTAAATGTTGAG 926
Qy 302 AACACCAAGTGTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTTG 361
Db 925 ATCGAAACACAGATGTTTACATGTTAAACAGGACATAGTTTAAACAAATGAAGAGAG 866
Qy 362 ACAGTATCCTCGTCTACAGAGATATCAACAGAAATCCTTTGATTTTAAAAACAAAT 421
Db 865 ACAAAATTATCCAGTAACTGAATTTTGTAGAACAATAATGTAATATGAAGTAAAAA 806
Qy 422 TATTGAAGTCTATTATTGCTGGAACAGCTTTTCAG 457
Db 805 TTGCACATGCTCTTCTCGAAGAACAGTGTTTCAG 770

RESULT 27

US-09-248-796A-778
; Sequence 778, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 778
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-778

Query Match 3.9%; Score 43.8; DB 4; Length 744;
Best Local Similarity 46.5%; Pred. No. 0.094;
Matches 141; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy 359 TTGACAGTATCCTGCTACAGAAGTATCAACAGAAATCCCTTTGATTTTAAAAACAA 418
Db 320 TAGATCAAAACCCACCTGAAGAAATTAGAAATATCTGATAATATTAAAGAAATTTAAAAACAA 379
Qy 419 AATTATTGAAGTCTATTATTGCTGAACAGAGCTTTTGCAGCCAGTTTACAGAAAGTCGTC 478
Db 380 AATTATTGGAGGAGTGGAGACGATATGAGATGAAATCATTTTGAAGATAGAGCTG 439
Qy 479 CTATTTTAAAGGAGTTCATATTGTTAAGTAATATGATAAGATTTTAAAGCAGTAGCGA 538
Db 440 ATGCTGCTGAAGAAGATGACGATGTAATACTAAACCTCTAGTCTAAACGCTGCAAGGAGAT 499
Qy 539 CTGACTCTCATCGTATGAGCCACGTTTAAATCACTTTTGGACAAATACTTTCAGCAGATTTTGA 598
Db 500 CAACCTGGTTCTAGGAAAGACAAAAGTCGTCACCTGTTCAAGATGCTTTAGAAATTTGGTG 559
Qy 599 TGGTAGTTCTTCCAAAGTAAATCTTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTG 658
Db 560 AACCAGAGATATAATAAATGATTAAGAATAGAAAGAAATTAATTGAAAGAAAG 619
Qy 659 AGA 661
Db 620 ATA 622

RESULT 28

US-08-956-171E-82
; Sequence 82, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 1598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-08-956-171E-82

Query Match          3.8%; Score 43.4; DB 4; Length 1598;
Best Local Similarity 51.3%; Pred. No. 0.36;
Matches 101; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 287 TTAAGAAATTGAACACACACAGTGTCTTAAACAGTGGTAAATCACAGATTACCTTAA 346
Db 5165 TGAAGAAATTAGTTCACGTCGAATTCACAACTAACATTTAAAAATCGGAGAGTAAATCG 5224

QY 347 AAGGAAAAGATGTTGACCAAGTATCCTCGTCTACAAGAGTATCAACAGAAAATCCTTTGA 406
Db 5225 AACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5284

QY 407 TTTTAAACAAATAATTATGAGTCTTATTGCTGTAACAGCTTTTGCAGCCAGTTTAC 466
Db 5285 TACTTAAACCGTGCACCACTTCATAGACTTGGTATTCAAGCAATTTGAACCACTTTAG 5344

QY 467 AAGAAAGTCGCTCTATT 483
Db 5345 TTGAAGTCGTCGCGATT 5361
```

```

RESULT 29
US-08-781-986A-82
; Sequence 82, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-82

Query Match          3.8%; Score 43.4; DB 4; Length 1598;
Best Local Similarity 51.3%; Pred. No. 0.36;
Matches 101; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```

QY 287 TTAAGAAATTGAACACACACAGTGTCTTAAACAGTGGTAAATCACAGATTACCTTAA 346
Db 5165 TGAAGAAATTAGTTCACGTCGAATTCACAACTAACATTTAAAAATCGGAGAGTAAATCG 5224

QY 347 AAGGAAAAGATGTTGACCAAGTATCCTCGTCTACAAGAGTATCAACAGAAAATCCTTTGA 406
Db 5225 AACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5284

QY 407 TTTTAAACAAATAATTATGAGTCTTATTGCTGTAACAGCTTTTGCAGCCAGTTTAC 466
Db 5285 TACTTAAACCGTGCACCACTTCATAGACTTGGTATTCAAGCAATTTGAACCACTTTAG 5344

QY 467 AAGAAAGTCGCTCTATT 483
Db 5345 TTGAAGTCGTCGCGATT 5361

RESULT 30
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```

```

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539) .. (855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619) .. (871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830) .. (1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1095846) .. (1095846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1119881) .. (1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1130881) .. (1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1310988) .. (1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1313224) .. (1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349473) .. (1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349491) .. (1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1470091) .. (1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1569020) .. (1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1602912) .. (1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1603734) .. (1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1637998) .. (1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1664854) .. (1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

Best Local Similarity 47.3%; Pred. No. 1.9;
Matches 165; Conservative 0; Mismatches

Qy	73	AAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAAGTCACCTTCTCAGAGTAAC	132
Db	1301470	AAAAGAAGCAGTTGGAATTAGCTAAGCAGCGAAAGAAACAGTAAAAAGTTGAGGATAT	1301529
Qy	133	TTAACAGGGTCTAACCGTCAAAATATCAAAATGAAAAACACTATTCCTGTAAGTAATAGAAAT	132
Db	1301530	AAAAATTGGCTTTGAAGAAATAAAATTTATTTTTAAAAATTTTTAAATTTTATTATTAATAT	1301589
Qy	193	GCATGGTTTGCATAATTACCTCTCCAGGAGCTATTTTATTAGAACCTAGTTTTTTTAA--TT	249
Db	1301590	TTTACTTCCTTCCAAACCTTTAGAAATCTTTGTTATTTTAAATTTATTCATTAATTTA	1301649
Qy	250	AATATTAATTCAGTTTGGCAGATATATAGTATAAAATGTTAAAGAAATGGAACACCAAC	309
Db	1301650	TATATTTATTTAAACCGGTGGAAATATGCAACAAAGAAATAAAAAATATAAAATATCGGTA	1301709
Qy	310	GTTGTTTAAACCAAGTCGTAAATCAGAGATATACCTTAAAAAGGAAAAAGATGTTGACCAAGTAT	369


```

; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

Db      1301530 AAAATTGGCTTTCAAGAAAAATAAAATTTATTTTAAAAATTTTAAATTTTATTTATTTAATAT 1301589
Qy      1301590 TTTACTTCCTCCAAACATTAGAAATCTTTGGTATTTTAAATTTATTTATTTATTTA---TT 249
Db      1301590 TTTACTTCCTCCAAACATTAGAAATCTTTGGTATTTTAAATTTATTTATTTATTTAATTTA 1301649
Qy      250 AATATTATTTCAAGTTTGCAGATATTTAGTATATAATTTGTTAAAGAAATTTGAACAACACCAA 309
Db      1301650 TATATTATTTAAACCGTGGAAATATGCACAAAGAAATAAAAAATATAAAATATATGCGGA 1301709
Qy      310 GTTGTTTTAACCAAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAGATCTTGACCAGTAT 369
Db      1301710 GTTACTGTAGTATAGATATAATGATTTTAAATTAAGGGAAGAGATAGATGATAAA 1301769
Qy      370 CCTCGTCTACAAGAAGTATCAACAGAAAAATCCCTTTGATTTTAAAAACAA 418
Db      1301770 TCTGGAATATTTAAAAAAGAACTAAACGCTAAAGTATATACATAA 1301818

RESULT 32
US-10-204-708-13/c
; Sequence 13, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 13
; LENGTH: 6113
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-13

Query Match 3.8%; Score 42.8; DB 4; Length 6113;
Best Local Similarity 45.9%; Pred. No. 0.36;
Matches 146; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy      5 TTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGTGCTA 64
Db      1024 TTCTATTTACTATTAATAATTTTAAATTCCTATTTTAAATACATAAAATTAATTA 965
Qy      65 TTAGCACTAAAAATGCCATTCCTATTTCTTTTCATCAATAAAAAATTTGAAGTCACCTTACAG 124
Db      964 TTTAAAAATTATATAAATTTCTATAAATTTAAACACATATAATTAATATAACCAACACCA 905
Qy      125 GAGTAACCTTTAACAGGGTCTAAGCGTCAATATCAATTTGAAAACACATATTCCTGTAAGTA 184
Db      904 AAATCAAAATACAAAAACAATTCGGCTCACCCCTAAATAACTCCCTCATACTACCTTTTATA 845
Qy      185 ATGAAATGCTGGTTTGCTAATTAATTCCTCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTT 244
Db      844 ATTACAATCTCTCCTCATCCCTAACCTTCAACACCACTAATTTATTCTCTACCATTTATAC 785
Qy      245 TTATTAAATATTATTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAAATTTGAACAAC 304

```

Db 784 TTTTTCATTTAAACTATCCACATATATATCTTATACCATATAAACCCTTTTAAACCT 725
 QY 305 ACCAAGTGTGTTTAAACCA 322
 Db 724 AACTTCTTCTATCAACA 707

RESULT 33
 US-09-539-333D-208/c
 ; Sequence 208, Application US/09539333D
 ; Patent No. 6476208
 ; GENERAL INFORMATION:
 ; APPLICANT: Cohen, Daniel
 ; APPLICANT: Blumenfeld, Marta
 ; APPLICANT: Chumakov, Ilya
 ; APPLICANT: Bouqueleret, Lydie
 ; APPLICANT: Bihain, Bernard
 ; APPLICANT: Essioux, Laurent
 ; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
 ; FILE REFERENCE: GENSET.047AUS
 ; CURRENT APPLICATION NUMBER: US/09/539,333D
 ; CURRENT FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: US 60/126,903
 ; PRIOR FILING DATE: 1999-03-30
 ; PRIOR APPLICATION NUMBER: US 60/131,971
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: US 60/132,065
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: US 60/143,928
 ; PRIOR FILING DATE: 1999-07-14
 ; PRIOR APPLICATION NUMBER: US 60/145,915
 ; PRIOR FILING DATE: 1999-07-27
 ; PRIOR APPLICATION NUMBER: US 60/146,453
 ; PRIOR FILING DATE: 1999-07-29
 ; PRIOR APPLICATION NUMBER: US 60/146,452
 ; PRIOR FILING DATE: 1999-07-29
 ; PRIOR APPLICATION NUMBER: US 60/162,288
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: US 09/416,384
 ; PRIOR FILING DATE: 1999-10-12
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 208
 ; LENGTH: 3001
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 1501
 ; OTHER INFORMATION: 99-26772-268 : polymorphic base C or T
 ; FEATURE:
 ; NAME/KEY: misc_binding
 ; LOCATION: 1481..1500
 ; OTHER INFORMATION: 99-26772-268.misl,
 ; FEATURE:
 ; NAME/KEY: misc_binding
 ; LOCATION: 1502..1520
 ; OTHER INFORMATION: 99-26772-268.mis2, complement
 ; FEATURE:
 ; NAME/KEY: primer_bind
 ; LOCATION: 1235..1254
 ; OTHER INFORMATION: upstream amplification primer
 ; FEATURE:
 ; NAME/KEY: primer_bind
 ; LOCATION: 1702..1722
 ; OTHER INFORMATION: downstream amplification primer, complement
 ; FEATURE:
 ; NAME/KEY: misc_binding
 ; LOCATION: 1489..1513
 ; OTHER INFORMATION: 99-26772-268 probe
 ; US-09-539-333D-208

Query Match 3.7%; Score 42.4; DB 4; Length 3001;
 Best Local Similarity 55.6%; Pred. No. 0.35;
 Matches 79; Conservative 1; Mismatches 62; Indels 0; Gaps 0;
 QY 311 TTGTTTTAACCAAGTGGTAAATCAGAGATTACTTAAAGGAAAGATGTTGACCAAGTATC 370
 Db 1585 TGGTTTTCAAATATGGTAATAGAGAAGTAATCTACACAGGGTAAAAAATATACCTAAAT 1526
 QY 371 CTCGTCTACAAGAGTATCAACAGAAATCTTTGATTTTAAACAACAAATTTGAAGT 430
 Db 1525 GAAATATTTTGAATTTTAACTATRAAATCTTGGATTTTACAAATTAATTTTAG 1466
 QY 431 CTATTATTGCTCAAAACAGCTTT 452
 Db 1465 TTGATTTTCTGTAGCATTTCT 1444
 RESULT 34
 US-09-248-796A-6131
 ; Sequence 6131, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 6131
 ; LENGTH: 2919
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-6131
 Query Match 3.7%; Score 42.2; DB 4; Length 2919;
 Best Local Similarity 42.6%; Pred. No. 0.39; 298; Indels 0; Gaps 0;
 Matches 221; Conservative 0; Mismatches 298; Indels 0; Gaps 0;
 QY 9 ATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATACAACTAAACGTCCTATTAG 68
 Db 2289 ATTATTACAATTTGATAAAATTTTCGATAAAGATTTTACTGAAACAAATTTGAAATATT 2348
 QY 69 CACTAAAAATGCCATTCCTATCTTTTCATCAATAAAAATTTGAAGTCATCTTACAGAGT 128
 Db 2349 TGCTGATATTGATAAAGCAAGTTTAAACCATAATTTGATGATATTTTAAATATTGAAAAAT 2408
 QY 129 AACTTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACATATTCCTGTAAGTAATGA 188
 Db 2409 ATATATTGCTCAAACTGTCAATTAATTTTCCATACCACTGGTCCAAAGCCAAAGC 2468
 QY 189 AAATGCTGGTTGCTAATTTACCTCTCCAGGAGCTATTTTATTAGAGCTAGTTTTTTTAT 248
 Db 2469 CAAAGGAACTAAATATTAGTTATATCTTCAACCACAAAGTTTAAAAAGTTGATGTTGTTA 2528
 QY 249 TAATATTATTTCAGTTTGGCAGATATTAGTATAATTTGTTAAAGAAATTTGAACACCA 308
 Db 2529 TGAGATTTTAATTCATTTTGTATAAAGTTTATCAATTAAGATCAAAACCATTAATGAAGA 2588
 QY 309 AGTTGTTTTAACCAAGTGGTAAATCAGAGATTACCTTAAAGGAAAGAGATGTTGACCAGTA 368
 Db 2589 AGTATTGTGCACATTTAATGAATTAATGAATTTTATGAAATATTCTTGATAA 2648
 QY 369 TCCTCGTCTACAAGAGTATCAACAGAAATTCCTTTGATTTTAAAAACAAATTTATTGAA 428
 Db 2649 TACTCGTCAAAATCTTTTAAAACTGTCATTAGCTTATGTTAAATTTAAATTTAGATGCAA 2708
 QY 429 GTCTATTATTGCTGAACACAGCTTTTGACGACGCTTTTACAGAAAGTCTCTCTATTAAAC 488


```
Db 2709 TTTTATTTTACTTGTTGTAATGAAATCAAACTTTTACAAATTAATGATCAAGTTATAA 2768
QY 489 AGAGTTCATATGTTAAGTAATCAATAAGATTTTAA 527
Db 2769 AATTTTACAAATTTTATTAAATGAAATGCATAATAAAA 2807

RESULT 35
US-09-248-796A-5237
; Sequence 5237, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5237
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5237

Query Match 3.7%; Score 41.6; DB 4; Length 1782;
Best Local Similarity 46.6%; Pred. No. 0.46; Mismatches 189; Indels 1; Gaps 1;
Matches 166; Conservative 0;

QY 72 TAAAAATGCGATTCCTATTCTTTTCATCAATAAAAAATGGAAGTCACTTCTACAGGAGTAAC 131
Db 831 TCAAGAAATTTACTCCAATTTCTTGACCTTAATAATGTTAATGTCATCACTAATGTTTGG 890
QY 132 TTTACAGGGTCAACGGTCAAAATCAATGGAACACACTATTCCTGTAAGTAATGAAAA 191
Db 891 TGGTACAAGTGGTGTGTGATGAGAACCAATACCACAAATAATGATATTTCTTCAAAA 950
QY 192 TGCTGGTTTGCTAAATACCTCTCCAGGAGCTATTTTATTAGAAGCTAGTTTATTATAA 251
Db 951 TCAGAAACATTAATAATCTGGCAATTTATGTTGGATATCTCGAAATTTTATCAATTGA 1010
QY 252 TATTATTTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAAATTGA-ACAACACCAAG 310
Db 1011 AATCAATTAATTTTTCAAAAGTTTGGTAGAATTTTAGAAAAATTTGATACATCAAC 1070
QY 311 TTGTTTTTACCAAGTGGTAAATCAGAGATTACCTTAAAGGAAAAGATGTTGACCAATC 370
Db 1071 CATTCGTTTCAAGTGGGAGTACTAACAAATAATAATCAATTAAGAAAATTTTAATAAATTT 1130
QY 371 CTCGCTCAACAAGTATCAACAGAAAATCCTTTGATTTTAAACAAAAATTTATG 426
Db 1131 GGCTGCCAACAGCAATAACAACAACAACAACAACAATACTATCACTAAACAGATTG 1186

RESULT 36
US-09-893-600-1/c
; Sequence 1, Application US/09893600
; Patent No. 6759205
; GENERAL INFORMATION:
; APPLICANT: Hansen, J. No. 6759205man
; TITLE OF INVENTION: Construction of a Strain of Bacillus subtilis 168 that Displays B
; FILE OF INVENTION: Sublancin Lantibiotic on the Surface of the Cell
; FILE REFERENCE: 108172-00057
; CURRENT APPLICATION NUMBER: US/09/893,600
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,449
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 1
; LENGTH: 2517
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The EcoRI-HindIII insert of the pLPVc integrative plasmid.
US-09-893-600-1

Query Match 3.7%; Score 41.6; DB 4; Length 2517;
Best Local Similarity 43.4%; Pred. No. 0.52;
Matches 191; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

QY 242 TTTTATTAATTAATTAATTTCAAGTTTGCAGATATAGTATATAATCTTAAAGAAATTTGAC 301
Db 1787 TTTTGGTTTTTCGAGTTCTCTCGAGTTTAACCTTCTTTAAATAGCTTTTCCATTTGTAAAAAC 1728
QY 302 AACACCAAGTTGTTTTTAACCAAGTGGTAAATCAGAGATTACCTTAAAGGAAAAGATGTTG 361
Db 1727 CTCGCCATTTGTTTGCATCTCGGATTACATTTTAAATATAAGGAAGATAATCTAAT 1668
QY 362 ACCAGTATCTCGTCTACAAGAAGTATCAACAGAAAAATCTTTGATTTTAAAAACAAAAT 421
Db 1667 TCAACCATATTTTATTTTATTTTGTAAAAATTAAGGAAAACCTTGTTCATAAATCATAGA 1608
QY 422 TATTGAAGTCTATTATGCTGAAAACAGCTTTTTCAGCCAGTTTACAAGAAAGTCGTCCTA 481
Db 1607 AAAAAAACTTTATTACCTTATCAATAAACAAGATTTGTTGTAATACACGAATCCCTTCTTC 1548
QY 482 TTTTAAACAGGAGTTTCATATTGTTTAACTAATATATAAGATTTTAAACAGTAGGAGCTG 541
Db 1547 AACTAACGGGCGAGGTAGTACATATAGAAAACCGACTGTAAAAAGTACAGTCGCAATTA 1488
QY 542 ACTCTCATCTGATGAGCAACGTTTAAATCACTTTGGACAATACTTTCAGCAGATTTGATGG 601
Db 1487 TCTCATATTATAAAGCCAGTCATTAGGCTATCTGACAATTCCTGAAATAGAGTTTCATAA 1428
QY 602 TAGTCTTCCAAAGTAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATATTGAGA 661
Db 1427 ACNAATCTCGCATGATACCAATCAACAACAAGATGATGATCTGTTAAAGATAGCGGTAAT 1368
QY 662 CCGTTGAGGTATTTTCTCA 681
Db 1367 ATATTGAATTAACCTTTATTA 1348

RESULT 37
US-09-248-796A-11532
; Sequence 11532, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 11532
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-11532

Query Match 3.7%; Score 41.4; DB 4; Length 603;
Best Local Similarity 43.8%; Pred. No. 0.35;
Matches 180; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 159 AATTGAAACACTATTCTCTGTAAGTAATGAAATGCTGTTGCTTAATTAACCTCTCCAGG 218
```

Db 114 AAGTCCCATCCAGTCTTGATAAAATGTGATTAAATATTTCTTGATACCCATCCAA 173
Qy 219 AGCTATTTTATAGAACTAGTTTTTTTATTAATAATATTTCAAGTTGGCAGATATTAG 278
Db 174 ATGTGATTATGATCAAAATATGTCTATTTGTAATAATGAATGATTAATATATGAGGAA 233
Qy 279 TATAAATGTTTAAAGAAATTCGAACAAACACCAAGTTGTTTTAACAGGTGGTAAATCAGAT 338
Db 234 AAATGATACAAAGGAAATGATATGGGACAGAAATTAAGAGAGAGATTTAAGAGAGAG 293
Qy 339 TACCTTAAAGGAAAGATGTTGACAGATATCTCGTCTACAAGAAAGTATCAACAGAAA 398
Db 294 TGACTTAAGTGAATATGATATACAAACAAATGAGTTAAGCGAAAGGATTCAGTGA 353
Qy 399 TCCTTTCATTTTAAACAAATATTTGAAGTCTATTTATGCTGTAACAGACTTTTGCAGC 458
Db 354 TGATTTAAGAAATAGCAACAAAGAACTGTTAGTGAATCTTTTGAAGAAATCAATGCA 413
Qy 459 CAGTTTACAAGAAAGTCTCTATTTTAAACAGGAGTTTATATGTTAAGTAATCATAA 518
Db 414 GAATTTGTTTGGACTTGGTCATGATTCATATGAATGTTCACTGGCATTTAGAAACATCA 473
Qy 519 AGATTTTAAAGCAGTAGCAGTCACTCTCATCTGATGAGCCACGTTTAAAT 569
Db 474 GTATATCCAGATTTATTTCTAGACTTCAGAGTTTTCGTGGAATAGAAAT 524

RESULT 38

US-09-248-796A-444
; Sequence 444, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 444
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-444

Query Match 3.7%; Score 41.4; DB 4; Length 825;
Best Local Similarity 44.1%; Pred. No. 0.39;
Matches 174; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

Qy 276 TAGTATAATGTTAAAGAAATTTGAACAAACCAAGTTGTTTAAACCAAGTGGTAAATCAGA 335
Db 132 TAGTCAACACAGAAAGATATACTGATCCAAACAACTAATAACCATATATAAAAAACA 191
Qy 336 GATTACCTTAAAGGAAAGATGTTGACCATGCTCTGCTACAAGAAAGTATCAACAGA 395
Db 192 AGGTAATTTGATATCAAGACGACTCTTATTAGATAATTTCAACAACTACAGAAACTTA 251
Qy 396 AAATCCTTTGATTTTAAACAAATATTTCAAGTCTATTTTCTGCTGAACAGCTTTTGC 455
Db 252 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 311
Qy 456 AGCCAGTTTACAAGAAAGTCTCTTATTTTAAACAGGAGTTTCAATATGTTAAGTAATCA 515
Db 312 TATATTAATGAAGAAACAGGGGAAATGCGAGCATTAATACAAGGAAATGACGACAAA 371
Qy 516 TAAAGATTTTAAAGCAGTAGCAGTCACTCTCATCTGATGAGCCACGTTTAAATCACTTT 575
Db 372 TAATGATAATAATGAAATTTATTAATATAGTTGATAAAGATATTCAGATAAATATTGA 431

Qy 576 GGACAATACTTCCAGCAGATTTGATGTTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTC 635
Db 432 TAGTTTTGAATTTTCAATAATTTAGTTTAAATAATGATTTAATGATATTAACAAGACTTTT 491
Qy 636 AGCAGTATTTACAGATGATATTTGAGACCGTTTGAGG 670
Db 492 AGGTATAAGTATGAAGAAATTTGCTAAATTTAAAG 526

RESULT 39

US-09-806-708B-23/c
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAE1 promoters
US-09-806-708B-23

Query Match 3.7%; Score 41.4; DB 4; Length 1055;
Best Local Similarity 20.4%; Pred. No. 0.43;
Matches 142; Conservative 166; Mismatches 382; Indels 5; Gaps 2;

Qy 223 ATTTTATTAGAGCTAGTTTTTTTATTAATAATATTTTCAAGTTGGCAGATATTAGTATA 282
Db 759 AYTWTGTCNNNNNNNNNNNTWCYTTTANAYKCCANNAANNAAGTCNNNTAYAAATTTT 700
Qy 283 AATGTTAAAGAAATTCGAACAAACCAAGTTGTTT-TTAAACCAAGTGGTAAATTCAGAGATTAC 341
Db 699 TRCTKTAANTATYTRWATTAKTTTTTRWATTTSTKNTNNNNNNAWACNNNNNNNAWKAITAM 640
Qy 342 CTTAAAGGAAAGATGTTGACCAAGTATCTCTGCTACAAGAAAGTATCAACAGAAATCC 401
Db 639 ATNWGCATNNNTWARATNNNTTRAGRGAAATNTKTGTTSTTAGWYTGATTGNNNNNNNNNC 580
Qy 402 TTTGATTTTAAAAACAAATATTTGAAGTCTATTATTGCTGAACAGCTTTTGCAGCCAG 461
Db 579 MTASCCCTGCTACTAATATGCGATSAKKRTATTTCTYRRRWANGWRTWTANMRCWRT 520
Qy 462 TTTTCAAGAAAGTCGCTCTATTTTAAACAGGAGTTTCATATTGTATTAAAGTAATCAATAAGA 521
Db 519 YTRTYCKSTAMS----CTWGNAMWYAAVSAGNTSSMARWTTANNTAAGYVRAAWAGTM 464
Qy 522 TTTTAAAGCAGTAGCAGCTGCTCATCGTATGAGCCAACTTTTATCCTTTTATCCTTTTGACAA 581
Db 463 WAAMANNNTTRYIAWTRWARYTTRYRSACNAKSMWRGATWRAATTTTAANNKAGAM 404
Qy 582 TACTTCAGCAGATTTGATGTTAGTTCTTCCAAAGTAAATCTTTTGAAGAGATTTTTCAGCACT 641
Db 403 WMTTAAAGNNNTWAAAAATKMAAWCARAYCCNNNNWAAACMAKWKWAWTWKYWGAACN 344
Qy 642 ATTACAGATGATATTGAGACCGTTGAGGTATTTTCTCACCAAGCCAAATCTTGTTCAG 701
Db 343 NNNKTAMYCRRAWMYSAWTTTTWAAATTSWKWYTTTTTRKTTWAAAAANNNNNNNNAKCKTT 284
Qy 702 AAGTGAACACATTTCTTTTATACAGCGCTCTTACAGGAAATTTATCCGATACAGCCG 761
Db 283 SAWWAWMMATWCTCGARTWGGATYMAAACCTTAAGAGNSMTYWCWAAATYMGTTMTNNNN 224

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 10:52:06 ; Search time 437 Seconds
(without alignments)
136222.067 Million cell updates/sec

Title: US-10-048-071-27
Perfect score: 1134
Sequence: 1 atgattcaatttcattaa.....ttaccacagtagcaacaac 1134

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1134	100.0	1134	4 AAF54744	Aaf54744 Nucleotid
2	1126	99.3	1134	6 ABN68601	Abn68601 Streptoco
3	1126	99.3	1137	6 ABX83362	Abx83362 Streptoco
5	1124.4	99.2	1137	8 ACA50247	ACA50247 Prokaryot
6	849	74.9	3200	2 AAV58285	AAV58285 S. pyogen
7	690.8	60.9	1134	8 ACA47964	ACA47964 Prokaryot
8	641.2	56.5	1124	6 ABN68600	Abn68600 Streptoco
9	641.2	56.5	55561	6 ABN71527_21	Continuation (22 o
10	608.2	53.6	1134	10 ABX08064	Abx08064 S. pneumo
11	608.2	53.6	21338	2 AAV52153	AAV52153 Streptoco
12	608.2	53.6	110000	10 ABS56454_20	Continuation (21 o
13	603.4	53.2	1137	4 AAS55711	Aas55711 Streptoco
14	603.4	53.2	1137	8 ACA49628	ACA49628 Prokaryot
15	408.2	36.0	110000	6 ABX90521_00	ABX90521 Genomic s
16	341.2	30.1	1131	4 AAS53173	Aas53173 Enterococ
17	338	29.8	1128	8 ACA32930	ACA32930 Prokaryot
18	330.6	29.2	8001	2 AAX13559	Aax13559 Enterococ
19	330.6	29.2	8001	6 ABS99354	Abs99354 Enterococ
20	317	28.0	1155	10 ADC91275	Adc91275 E. faeciu
21	315.6	27.8	1128	8 ACA33953	ACA33953 Prokaryot

22	247.4	21.8	1131	8 ACA46779	ACA46779 Prokaryot
23	247.4	21.8	1158	6 ABN92866	Abn92866 Staphyloc
24	243.4	21.5	1131	8 ACA47854	ACA47854 Prokaryot
25	242	21.3	110000	10 ADF77343_00	Adf77343 Lactic ac
26	236.2	20.8	1131	8 ACF74365	Acf74365 Staphyloc
27	236.2	20.8	1134	4 AAS54991	Aas54991 Staphyloc
28	236.2	20.8	1134	6 ABN86787	Abn86787 Staphyloc
29	236.2	20.8	1134	10 ADD26271	Add26271 Staphyloc
30	236.2	20.8	1134	10 ADD26273	Add26273 Staphyloc
31	236.2	20.8	2347	2 AAV74464	Aav74464 Staphyloc
32	233	20.5	1134	2 AAZ31005	Aaz31005 Partial d
33	233	20.5	1134	4 AAS54965	Aas54965 Staphyloc
34	233	20.5	1134	4 AAF54735	Aaf54735 Nucleotid
35	233	20.5	1134	8 ACA20030	ACA20030 Prokaryot
36	233	20.5	1134	10 AAD62920	Aad62920 Staphyloc
37	230	20.3	1128	6 AAS52032	Aas52032 Staphyloc
38	200.4	17.7	4736	6 ABQ70950	Abq70950 Listeria
39	198.8	17.5	110000	6 ABQ67196_2	Continuation (3 of
40	198.8	17.5	110000	6 ABQ69245_00	ABQ69245 Listeria
41	186	16.4	110000	6 ABA03041_00	ABA03041 Listeria
42	185.6	16.4	1146	8 ACA36279	ACA36279 Prokaryot
43	181.2	16.0	1218	9 ADB07181	ADB07181 Alloiococ
44	181.2	16.0	1218	9 ADB07183	ADB07183 Alloiococ
45	181.2	16.0	1218	9 ADB07179	ADB07179 Alloiococ

ALIGNMENTS

RESULT 1
AAF54744
ID AAF54744 standard; DNA; 1134 BP.
XX
AC AAF54744;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a partial dnaN (beta subunit) polypeptide.
XX
KW dnaB; Gram positive bacteria; polC; dnaE; hola; holB; dnaX; ssb;
KW dnaG; dnaB; antibiotic; replication; cell growth; cell death;
KW bacterial infection; ss.
XX
OS Streptococcus pyogenes.
XX
PN WO200109164-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US020666.
XX
PR 29-JUL-1999; 99US-0146178P.
XX
PA (UTRQ) UNIV ROCKEFELLER.
XX
PI O'donnell ME, Bruck I, Zhang D, Whipple R;
XX
DR WPI; 2001-147453/15.
XX
P-PSDB; AAB31944.

Isolated DNA molecule from a Gram positive bacterium encoding DNA replication proteins used to identify compounds which have antibiotic activity.

Claim 38; Page 61-62; 239pp; English.

The present sequence encodes a partial dnaN polypeptide. The specification describes DNA molecules from Gram positive bacteria, which comprise a coding region from a polC, dnaE, hola, holB, dnaX, ssb, dnaG or a dnaB gene. These sequences encode proteins that replicate the chromosome of Gram positive bacteria. They are used for sequencing and amplification of DNA and in drug discovery to identify compounds which have antibiotic activity through interference with replication. They are

CC used in methods for identifying compounds that are active at the level of
CC DNA replication and result in arrest of cell growth or cell death of
CC bacteria to treat bacterial infections in animals
XX

SQ Sequence 1134 BP; 383 A; 201 C; 178 G; 372 T; 0 U; 0 Other;
Query Match 100.0%; Score 1134; DB 4; Length 1134;
Best Local Similarity 100.0%; Pred. No. 2.1e-232;
Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGATTCATATTTCAATTAATGCGACATTTATTCATGCTTTAAATACAACTAAACGT 60
DB 1 ATGATTCATATTTCAATTAATGCGACATTTATTCATGCTTTAAATACAACTAAACGT 60

QY 61 GCTATTAGCACTAAATGCGCATTCCTATTCCTCATCAATAAAATTTGAAGTCACTTCT 120
DB 61 GCTATTAGCACTAAATGCGCATTCCTATTCCTCATCAATAAAATTTGAAGTCACTTCT 120

QY 121 ACAGGAGTAACCTTTAACAGGGTCTAACGGTCAAATATCAATTTGAAAAACACTATTCCTGTA 180
DB 121 ACAGGAGTAACCTTTAACAGGGTCTAACGGTCAAATATCAATTTGAAAAACACTATTCCTGTA 180

QY 181 AGTAATGAAATGCTGGTGTGCTAATTAACCTTCAGGAGCTATTTTATTAGAACCTAGT 240
DB 181 AGTAATGAAATGCTGGTGTGCTAATTAACCTTCAGGAGCTATTTTATTAGAACCTAGT 240

QY 241 TTTTATTAATTAATTTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATTCGA 300
DB 241 TTTTATTAATTAATTTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATTCGA 300

QY 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAAAGGAAAGATGTT 360
DB 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAAAGGAAAGATGTT 360

QY 361 GACCAAGTATCCTCGTCTACAGAAATGATCAACAGAAATCTCTTGTATTTAAACCAAAA 420
DB 361 GACCAAGTATCCTCGTCTACAGAAATGATCAACAGAAATCTCTTGTATTTAAACCAAAA 420

QY 421 TTATTTGAAGTCTATTATTTGCTGAAACAGCTTTTGCAGCAGTTTACAAAGAAAGTGGTCT 480
DB 421 TTATTTGAAGTCTATTATTTGCTGAAACAGCTTTTGCAGCAGTTTACAAAGAAAGTGGTCT 480

QY 481 ATTTTAAACAGGAGTTTCAATATTCATATTAAGTAAATCAATAAAGATTTTAAAGCAGTAGCGACT 540
DB 481 ATTTTAAACAGGAGTTTCAATATTCATATTAAGTAAATCAATAAAGATTTTAAAGCAGTAGCGACT 540

QY 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTTGGACATACCTTCAGCAGATTTGATG 600
DB 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTTGGACATACCTTCAGCAGATTTGATG 600

QY 601 GPAGTTCTTCCAGTAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTCAG 660
DB 601 GPAGTTCTTCCAGTAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTCAG 660

QY 661 ACCGTGTAGGTATTTTCTCAACAGCCAAATCTTTGTTGAGAAAGTGAACACATTTCTTTT 720
DB 661 ACCGTGTAGGTATTTTCTCAACAGCCAAATCTTTGTTGAGAAAGTGAACACATTTCTTTT 720

QY 721 TATACAGCCTCTTAAAGGAATTTATCCGATACAGCCGTTTATTAATGACAGAAATTT 780
DB 721 TATACAGCCTCTTAAAGGAATTTATCCGATACAGCCGTTTATTAATGACAGAAATTT 780

QY 781 GAGAGGAGGTGTTTTCATACCAATCCCTTCGCCAGCTATGGAAGTGCCTTCTTGT 840
DB 781 GAGAGGAGGTGTTTTCATACCAATCCCTTCGCCAGCTATGGAAGTGCCTTCTTGT 840

QY 841 ATTCTTAATGCTACTCAAAATGCTATGTTAAGTTGAGATTAATCAAAATCATATTTCA 900
DB 841 ATTCTTAATGCTACTCAAAATGCTATGTTAAGTTGAGATTAATCAAAATCATATTTCA 900

QY 901 GCTCATGTTAACTCACTCAGGTGTTAGGTAAGGTAAGGATTTAGATATTTAGTACAG 960
DB 901 GCTCATGTTAACTCACTCAGGTGTTAGGTAAGGTAAGGATTTAGATATTTAGTACAG 960
```

```
QY 961 TCTGTTAGTGAATTAACCTATCAGCTTCAATCAACTTACCTTATTGAGTCTTTAAAGCT 1020
DB 961 TCTGTTAGTGAATTAACCTATCAGCTTCAATCAACTTACCTTATTGAGTCTTTAAAGCT 1020

QY 1021 ATTAAAAAGTGAACAGTAATAAATTCATTTCTTATCAGCAGTTTCAGCATTACCTTAACA 1080
DB 1021 ATTAAAAAGTGAACAGTAATAAATTCATTTCTTATCAGCAGTTTCAGCATTACCTTAACA 1080

QY 1081 CCAGCGATGAGGAAGAAAGTTTATCAATTAATTAACACCAAGTACGAAACAAAC 1134
DB 1081 CCAGCGATGAGGAAGAAAGTTTATCAATTAATTAACACCAAGTACGAAACAAAC 1134

RESULT 2
ABN68601
ID ABN68601 standard; DNA; 1134 BP.
XX
AC ABN68601;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 5115.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;
PI Tettelin H;
XX
WPI: 2002-352536/38.
P-PSDB; ABP27970.
XX
New Streptococcus protein for the treatment or prevention of infection or
disease caused by Streptococcus bacteria, such as meningitis, and for
detecting a compound that binds to the protein.
XX
Claim 7; Page 3675; 4525pp; English.
XX
The invention relates to a protein (ABP25413-ABP30895) from group B
streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
the specification. The proteins have antibacterial and antiinflammatory
activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
antibodies that bind (I) are used in the manufacture of medicaments for
the treatment or prevention of infection or disease caused by
Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
Nucleic acids encoding (I) is used to determine whether a compound binds to
biological sample. (I) is used to determine whether a compound binds to
used as a vaccine or diagnostic composition. The disease caused by
Streptococcus that is prevented or treated may be meningitis. Nucleic
acid encoding (I) may be used to recombinantly produce (I) and may be
used in gene therapy. Antibodies to (I) are used for affinity
chromatography, immunoassays, and distinguishing/identifying
Streptococcus proteins
XX
```

SQ Sequence 1134 BP; 384 A; 200 C; 177 G; 373 T; 0 U; 0 Other;

Query Match 99.3%; Score 1126; DB 6; Length 1134;
 Best Local Similarity 99.6%; Pred. No. 1.1e-230;
 Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAATAAGCT 60
 DB 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAATAAGCT 60

QY 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAGTCACTTCT 120
 DB 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAGTCACTTCT 120

QY 121 ACAGAGTAACCTTTTAAACGGGTCTAAACGGTCAAAATATCAATTTGAAAAACAATTCCTGTA 180
 DB 121 ACAGAGTAACCTTTTAAACGGGTCTAAACGGTCAAAATATCAATTTGAAAAACAATTCCTGTA 180

QY 181 ACTAATGAAGTCTGGTTGCTTAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 DB 181 ACTAATGAAGTCTGGTTGCTTAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240

QY 241 TTTTATTATTAATTTTCAAGTTTCCAGATATTTAGTATAAATGTTAAAGAAATTTGAA 300
 DB 241 TTTTATTATTAATTTTCAAGTTTCCAGATATTTAGTATAAATGTTAAAGAAATTTGAA 300

QY 301 CAACACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTAAAGGAAAAAGATGTT 360
 DB 301 CAACACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTAAAGGAAAAAGATGTT 360

QY 361 GACCAATATCTCGTCTACAGAGATCAACAGAAATCTTTGATTTTAAACAAAA 420
 DB 361 GACCAATATCTCGTCTACAGAGATCAACAGAAATCTTTGATTTTAAACAAAA 420

QY 421 TTATTGAAGTCTATTATTCTGAAACAGCTTTTGCAGCAGTTTCAAGAAAGTCTGCT 480
 DB 421 TTATTGAAGTCTATTATTCTGAAACAGCTTTTGCAGCAGTTTCAAGAAAGTCTGCT 480

QY 481 ATTTTAAACAGGAGTTTCAATTTGTTAGTAAATCAATAAAGATTTTAAAGCAGTAGCGACT 540
 DB 481 ATTTTAAACAGGAGTTTCAATTTGTTAGTAAATCAATAAAGATTTTAAAGCAGTAGCGACT 540

QY 541 GACTCTCATCGTATGAGCAAGTTTATCACTTTGGCAATACCTTCAGCAGATTGATG 600
 DB 541 GACTCTCATCGTATGAGCAAGTTTATCACTTTGGCAATACCTTCAGCAGATTGATG 600

QY 601 GTAGTTCTTCCAGTAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTGAG 660
 DB 601 GTAGTTCTTCCAGTAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTGAG 660

QY 661 ACCGTGAGGTATTTTCTCACCAGCCAAATCTTTGTTGAGAGTGAACACATTTCTTTT 720
 DB 661 ACCGTGAGGTATTTTCTCACCAGCCAAATCTTTGTTGAGAGTGAACACATTTCTTTT 720

QY 721 TATACAGCTCTTAGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
 DB 721 TATACAGCTCTTAGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780

QY 781 GAGACGAGGTGTTTTCATACCCAAATCCCTTTGCGCCAGCTATGGAAGTCCCTCTTTG 840
 DB 781 GAGACGAGGTGTTTTCATACCCAAATCCCTTTGCGCCAGCTATGGAAGTCCCTCTTTG 840

QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTTAAAGTATGAGTACTCAAAATCATATTTCA 900
 DB 841 ATTTCTAATGCTACTCAAAATGGTACTGTTTAAAGTATGAGTACTCAAAATCATATTTCA 900

QY 901 GCTCATGTTAACTCACCTGAGGTGTTGTAAGGTAACAGGATTTAGATTTGTTAGTCAG 960
 DB 901 GCTCATGTTAACTCACCTGAGGTGTTGTAAGGTAACAGGATTTAGATTTGTTAGTCAG 960

QY 961 TCTGGTAGTGATTTAACTATCACTTCAATCAACTTACCTTTTGGTCTTTTAAAGCT 1020
 DB 961 TCTGGTAGTGATTTAACTATCACTTCAATCAACTTACCTTTTGGTCTTTTAAAGCT 1020

QY 1021 ATTTAAAGTGAACAGTAAAAATTCATTTCTTATCAACAGTTGACCATTTCCCTTAACA 1080
 DB 1021 ATTTAAAGTGAACAGTAAAAATTCATTTCTTATCACCAGTTGACCATTTCCCTTAACA 1080

QY 1081 CCAGCGATAGGAGAAAGTTTATCCAAATTTATACACAGTACGACAAAC 1134
 DB 1081 CCAGCGATAGGAGAAAGTTTATCCAAATTTATACACAGTACGACAAAC 1134

RESULT 3
 ABK83362
 ID ABK83362 standard; DNA; 1137 BP.
 XX AC ABK83362;
 XX 12-AUG-2002 (first entry)
 XX Streptococcus pyogenes dnaN gene.
 XX DNA polymerase III holoenzyme subunit protein; replicase; polymerase;
 KW antibacterial; vaccine; gene; ds.
 XX Streptococcus pyogenes.
 XX WO200234936-A2.
 XX 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-US048396.
 XX 27-OCT-2000; 2000US-0244023P.
 XX (REPL-) REPLIDYNE INC.
 XX McHenry CS, Bullard JM, Janjic N, Manhardt EL, Kery V;
 PI Williams JC;
 XX WPI; 2002-463366/49.
 P-PSDB; ABG61623.
 Novel bacterial DNA polymerase III subunit proteins and genes encoding
 the proteins, for reconstituting replicases and polymerases for
 sequencing, amplification and screening compounds that modulate
 polymerase.
 Claim 10; Fig 21A; 268pp; English.
 The invention describes an isolated bacterial, preferably Streptococcus
 pyogenes DNA polymerase III holoenzyme subunit proteins (I). The protein
 and the polynucleotide encoding it are useful for reconstituting
 replicases and polymerases for sequencing, amplification and screening
 for compounds which modulate the function of polymerase or replicase. An
 antibody to the protein is useful as therapeutic compound to passively
 immunise an animal to protect the animal from bacteria susceptible to
 treatment by such antibodies, preferably Streptococcus pyogenes, as
 reagents in assays to detect infection by such bacteria, and/or as tools
 to screen expression libraries and/or to recover desired proteins from a
 mixture of protein and other contaminants. This sequence encodes a
 Streptococcus pyogenes polymerase III holoenzyme sub-unit
 Sequence 1137 BP; 386 A; 200 C; 177 G; 374 T; 0 U; 0 Other;

Query Match 99.3%; Score 1126; DB 6; Length 1137;
 Best Local Similarity 99.6%; Pred. No. 1.1e-230;
 Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAATAAGCT 60
 DB 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAATAAGCT 60

QY 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAGTCACTTCT 120
 DB 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAGTCACTTCT 120

Db 61 GCTATTAGCACTAAATAATGCAATTCCTATTCTTTCATCAATAAATAATGAAGTCACCTTCT 120
Qy 121 ACAGGAGTAACCTTTAAACAGGGCTTAACGGTCAATATCAATTAATGAACACACTATTCCTGTA 180
Db 121 ACAGGAGTAACCTTTAAACAGGGCTTAACGGTCAATATCAATTAATGAACACACTATTCCTGTA 180
Qy 181 AGTAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 AGTAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy 241 TTTTATTAATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
Db 241 TTTTATTAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
Qy 301 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGATTAACCTTAAAGGAAAGATGTT 360
Db 301 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGATTAACCTTAAAGGAAAGATGTT 360
Qy 361 GACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 GACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy 421 TTATTGAAGTCTATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 TTATTGAAGTCTATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 ATTTTAAACAGGAGTTCATATTTGTAATTAAGTAATCAATAAAGATTTTAAAGCAGTAGGACT 540
Db 481 ATTTTAAACAGGAGTTCATATTTGTAATTAAGTAATCAATAAAGATTTTAAAGCAGTAGGACT 540
Qy 541 GACTCTCATGCTATGAGCCAAAGCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 GACTCTCATGCTATGAGCCAAAGCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 601 GTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATATATGAG 660
Db 601 GTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATATATGAG 660
Qy 661 ACCGTTGAGTATTTTTCACCAAGCCAAATCTTGTTCAGAGTGAACACATTTCTTTT 720
Db 661 ACCGTTGAGTATTTTTCACCAAGCCAAATCTTGTTCAGAGTGAACACATTTCTTTT 720
Qy 721 TATACAGCGCTCTTGAAGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
Db 721 TATACAGCGCTCTTGAAGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
Qy 781 GAGACGAGGTTGTTTTCATACCAATCCCTTCGCGCTATGGAACGTCGCTTCTTG 840
Db 781 GAGACGAGGTTGTTTTCATACCAATCCCTTCGCGCTATGGAACGTCGCTTCTTG 840
Qy 841 ATTTCTTAATGCTACTCAAAATGCTGTTAGCTTGAATTAATCAAAATCAATTTCA 900
Db 841 ATTTCTTAATGCTACTCAAAATGCTGTTAGCTTGAATTAATCAAAATCAATTTCA 900
Qy 901 GCTCATGTTAACTCACTGAGGTTGCTGAGTAAAGGTAAGAGGATTTAGATTTGTTAGTCAG 960
Db 901 GCTCATGTTAACTCACTGAGGTTGCTGAGTAAAGGTAAGAGGATTTAGATTTGTTAGTCAG 960
Qy 961 TCTGTTAGTGAATTAATCAATGCTTCAATCCACTTACCTTATTTAGCTTTTAAAGCT 1020
Db 961 TCTGTTAGTGAATTAATCAATGCTTCAATCCACTTACCTTATTTAGCTTTTAAAGCT 1020
Qy 1021 ATTTAAAGTGAAACAGTAAATTTTCTTATCAACAGTTCACCAATTTACCTTCAAC 1080
Db 1021 ATTTAAAGTGAAACAGTAAATTTTCTTATCAACAGTTCACCAATTTACCTTCAAC 1080
Qy 1081 CCAGCGATGAGGAAAGATTTTATCAATTAATTAACAGTACCAACAAAC 1134
Db 1081 CCAGCGATGAGGAAAGATTTTATCAATTAATTAACAGTACCAACAAAC 1134

RESULT 4
ACA50247

ID ACA50247 standard; DNA; 1137 BP.

XX ACA50247;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #31904.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
drug design; gene.

XX Streptococcus pyogenes.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU46377.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 38117; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1137 BP; 386 A; 200 C; 177 G; 374 T; 0 U; 0 Other;

Query Match 99.3%; Score 1126; DB 8; Length 1137;
Best Local Similarity 99.6%; Pred. No. 1.1e-230;
Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 1 ATGATTCAATTTTCAATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGT 60

QY 61 GCTATTAGCACTAAAATGCCATTCCTATTTCTTTTATCATCAATAAAAATTTGAAGTCATCTCT 120
DB 61 GCTATTAGCACTAAAATGCCATTCCTATTTCTTTTATCATCAATAAAAATTTGAAGTCATCTCT 120

QY 121 ACAGGAGTAACTTTAACAGGCTTAAACGGTCAAAATCAATTAATGAAACACATTTCTCTGA 180
DB 121 ACAGGAGTAACTTTAACAGGCTTAAACGGTCAAAATCAATTAATGAAACACATTTCTCTGA 180

QY 181 AGTAATGAAATGCTGGTTGCTAAATACCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
DB 181 AGTAATGAAATGCTGGTTGCTAAATACCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 240

QY 241 TTTTATTAATTAATTTTCAAGTTTGCAGATATTTAGTATTAATGTTTAAAGAAATGAA 300
DB 241 TTTTATTAATTAATTTTCAAGTTTGCAGATATTTAGTATTAATGTTTAAAGAAATGAA 300

QY 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360
DB 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360

QY 361 GACCAAGTATCTCTGCTCAACAGAGTATCAACAGAAATCTCTTGCATTTTAAACCAAAA 420
DB 361 GACCAAGTATCTCTGCTCAACAGAGTATCAACAGAAATCTCTTGCATTTTAAACCAAAA 420

QY 421 TTATTTGAAGTCTATTTATTTGCTGAAAAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCT 480
DB 421 TTATTTGAAGTCTATTTATTTGCTGAAAAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCT 480

QY 481 ATTTTAAACAGGAGTTTATTTGATTTAGTATTAATCAATTAAGATTTTAAAGCAGTACGACT 540
DB 481 ATTTTAAACAGGAGTTTATTTGATTTAGTATTAATCAATTAAGATTTTAAAGCAGTACGACT 540

QY 541 GACTCTCATGATGAGCCAAAGTTTAAATCACTTTGGCAATACCTTACGACAGATTTGATG 600
DB 541 GACTCTCATGATGAGCCAAAGTTTAAATCACTTTGGCAATACCTTACGACAGATTTGATG 600

QY 601 GTAGTTCTTCCAGTAATTTTGTGAGAGAAATTTTCCGATACAGACCGTTTATTAATGACAGATT 660
DB 601 GTAGTTCTTCCAGTAATTTTGTGAGAGAAATTTTCCGATACAGACCGTTTATTAATGACAGATT 660

QY 661 ACCGTTGAGTATTTTCTCACCAGCCAAATCTTTGTCAGAGTGAACACATTTCTTTT 720
DB 661 ACCGTTGAGTATTTTCTCACCAGCCAAATCTTTGTCAGAGTGAACACATTTCTTTT 720

QY 721 TATACACGCTCTTGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
DB 721 TATACACGCTCTTGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780

QY 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCAGCTATGGAAGCTGCTTTCTTG 840
DB 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCAGCTATGGAAGCTGCTTTCTTG 840

QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAGCTTTGAGATTACTCAAAATCATATTTCA 900
DB 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAGCTTTGAGATTACTCAAAATCATATTTCA 900

QY 901 GCTCATGTTAACTCACTTCACTGAGTTGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 960
DB 901 GCTCATGTTAACTCACTTCACTGAGTTGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 960

QY 961 TCTGTTAGTGAATTAATCATGCTTCAATCCAACTTACCTTTATTTGAGTCTTTTAAAGCT 1020
DB 961 TCTGTTAGTGAATTAATCATGCTTCAATCCAACTTACCTTTATTTGAGTCTTTTAAAGCT 1020

QY 1021 ATTTAAAGTGAACAGTAAAAATTCATTTCTTATCACCAGTTTCGACCATTCACCCTAACA 1080

Db 1021 ATTTAAAGTGAACAGTAAAAATTCATTTCTTATCACCAGTTTCGACCATTCACCCTAACA 1080

QY 1081 CCAGCGGATGAGGAAGAAAGTTTATCCAAATTAATTAACACGATGACGAACAAAC 1134

Db 1081 CCAGCGGATGAGGAAGAAAGTTTATCCAAATTAATTAACACGATGACGAACAAAC 1134

RESULT 5
ABK83363
ID ABK83363 standard; DNA; 1137 BP.
XX
AC ABK83363;
XX
DT 12-AUG-2002 (first entry)
XX
DE DNA encoding Streptococcus pyogenes DnaN protein.
XX
KW DNA polymerase III holoenzyme subunit protein; replicase; polymerase;
KW antibacterial; vaccine; gene; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234936-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US048396.
XX
PR 27-OCT-2000; 2000US-0244023P.
XX
PA (REPL-) REPLIDYNE INC.
XX
PI McHenry CS, Bullard JM, Janjic N, Manhardt BL, Kery V;
PI Williams JC;
XX
WPI; 2002-463366/49.
DR P-PSDB; ABG61623.
XX
PT Novel bacterial DNA polymerase III subunit proteins and genes encoding
PT the proteins, for reconstituting replicases and polymerases for
PT sequencing, amplification and screening compounds that modulate
PT polymerase.
XX
PS Disclosure; Page 217-218; 268pp; English.
XX
CC The invention describes an isolated bacterial, preferably Streptococcus
CC pyogenes DNA polymerase III holoenzyme subunit proteins (I). The protein
CC and the polynucleotide encoding it are useful for reconstituting
CC replicases and polymerases for sequencing, amplification and screening
CC for compounds which modulate the function of polymerase or replicase. An
CC antibody to the protein is useful as therapeutic compound to passively
CC immunise an animal to protect the animal from bacteria susceptible to
CC treatment by such antibodies, preferably Streptococcus pyogenes, as
CC reagents in assays to detect infection by such bacteria, and/or as tools
CC to screen expression libraries and/or to recover desired proteins from a
CC mixture of protein and other contaminants. This sequence encodes a
CC Streptococcus pyogenes polymerase III holoenzyme sub-unit
XX
SQ Sequence 1137 BP; 387 A; 200 C; 177 G; 373 T; 0 U; 0 Other;

Query Match 99.2%; Score 1124.4; DB 6; Length 1137;
Best Local Similarity 99.5%; Pred. NO. 2.3e-230;
Matches 1128; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGT 60

DB 1 ATGATTCAATTTTCAATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGT 60

QY 61 GCTATTAGCACTAAAATGCCATTCCTATTTCTTTTATCATCAATAAAAATTTGAAGTCATCTCT 120

DB 61 GCTATTAGCACTAAAATGCCATTCCTATTTCTTTTATCATCAATAAAAATTTGAAGTCATCTCT 120

121 ACAGGAGTAATCTTTAAACAGGCTCTAACGGTCAATATCAATGAAACACACTATTCCTGTA 180
 121 ACAGGAGTAATCTTTAAACAGGCTCTAACGGTCAATATCAATGAAACACACTATTCCTGTA 180
 181 AGTAATGAAATGCTGTTGCTAAATTTACCTCTCCAGGAGCTATTTTATGAAAGCTAGT 240
 181 AGTAATGAAATGCTGTTGCTAAATTTACCTCTCCAGGAGCTATTTTATGAAAGCTAGT 240
 241 TTTTATTAATATATTTTCAAGTTTCCAGATATTAGTATAATGTTAAAGAAATGAA 300
 241 TTTTATTAATATATTTTCAAGTTTCCAGATATTAGTATAATGTTAAAGAAATGAA 300
 301 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 360
 301 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 360
 361 GACCAATATCTCTGCTACAGAGATGTAATCAAGAAATCTTTGATTTTAAACAA 420
 361 GACCAATATCTCTGCTACAGAGATGTAATCAAGAAATCTTTGATTTTAAACAA 420
 421 TTATTGAAGTCTATTTTCAAGTGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 480
 421 TTATTGAAGTCTATTTTCAAGTGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 480
 481 ATTTTAAACAGGAGTTCATATTTGATTAAGTAATCAATGAAAGATTTTAAAGCAGTAGGACT 540
 481 ATTTTAAACAGGAGTTCATATTTGATTAAGTAATCAATGAAAGATTTTAAAGCAGTAGGACT 540
 541 GACTCTCATGCTATGAGCAACGTTTAAATCACTTTGGCAATACCTTCAGCAGATTGATG 600
 541 GACTCTCATGCTATGAGCAACGTTTAAATCACTTTGGCAATACCTTCAGCAGATTGATG 600
 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGATGATATGAG 660
 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGATGATATGAG 660
 661 ACGTTGAGGTATTTTCTCCACAGCAATCTTTGTTGAGAGTGAACACATTTCTTTT 720
 661 ACGTTGAGGTATTTTCTCCACAGCAATCTTTGTTGAGAGTGAACACATTTCTTTT 720
 721 TATACAGGCTCTTGAAGGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
 721 TATACAGGCTCTTGAAGGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCCGACGCTATGGAACGTTGCTTCTTG 840
 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCCGACGCTATGGAACGTTGCTTCTTG 840
 841 ATTTCTAATGCTACTCAAAATGCTACTGTTAAGCTTGAGATTACTCAAAATCATATTTCA 900
 841 ATTTCTAATGCTACTCAAAATGCTACTGTTAAGCTTGAGATTACTCAAAATCATATTTCA 900
 901 GCTCATGTTAACTCACTGAGTTGGTAAAGTAAACAGGAGTTTATGATTTGTTAGTCAG 960
 901 GCTCATGTTAACTCACTGAGTTGGTAAAGTAAACAGGAGTTTATGATTTGTTAGTCAG 960
 961 TCTGCTAGTATTTAACTCATGCTTCAATCCAACTTACCTTTATGAGTCTTTAAAGCT 1020
 961 TCTGCTAGTATTTAACTCATGCTTCAATCCAACTTACCTTTATGAGTCTTTAAAGCT 1020
 1021 ATTTAAAGTGAACAGTAAAGTATTTTATCAATTTCTTATCACCAGTTTCCACCTTAACA 1080
 1021 ATTTAAAGTGAACAGTAAAGTATTTTATCAATTTCTTATCACCAGTTTCCACCTTAACA 1080
 1081 CCAGCGATGAGGAAAGTATTTTATCAATTTATACACAGTACGACAAAC 1134
 1081 CCAGCGATGAGGAAAGTATTTTATCAATTTATACACAGTACGACAAAC 1134

RESULT 6
 AAV58285
 ID AAV58285 standard; DNA; 3200 BP.
 XX

AC AAV58285;
 XX 18-DEC-1998 (first entry)
 XX S. pyogenes SP-7-44 genomic DNA.
 DE Probe; diagnosis; infection; detection; ds.
 XX Streptococcus pyogenes.
 XX WO9842845-A1.
 XX 01-OCT-1998.
 XX 23-MAR-1998; 98WO-JP001288.
 XX 25-MAR-1997; 97JP-00071077.
 XX (FUSO) FUSO PHARM IND LTD.
 XX Ueyama H, Abe K, Keshi H, Matsuhisa A;
 XX WPI; 1998-532009/45.
 XX New DNA probes, e.g. SP-6-28 or SP-7-44 - useful for, e.g. diagnosis of
 XX Streptococcus pyogenes infection.
 XX Claim 2; Page 19-21; 48pp; Japanese.
 XX AAV58284-V58289 are novel genomic DNA sequences which can be used as DNA
 XX probes for the diagnosis of Streptococcus pyogenes infection. These
 XX probes provide for simple and highly specific detection of S. pyogenes in
 XX biological samples such as blood
 XX Sequence 3200 BP; 1120 A; 547 C; 523 G; 1010 T; 0 U; 0 Other;
 Query Match 74.9%; Score 849; DB 2; Length 3200;
 Best Local Similarity 98.7%; Pred. No. 1.6e-171;
 Matches 866; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
 QY 1 ATGATTCAATTTTCAATTAATCGCAGATTTATTTATTCATGCTTTTAAATGCAACTTAAACGT 60
 Db 2324 ATGATTCAATTTTCAATTAATCGCAGATTTATTTATTCATGCTTTTAAATGCAACTTAAACGT 2383
 QY 61 GCTATTAGCACTAAATGCAATTCCTATTCTTTTCATCAATATAAATTTGAAGTCACTTCT 120
 Db 2384 GCTATTAGCACTAAATGCAATTCCTATTCTTTTCATCAATATAAATTTGAAGTCACTTCT 2443
 QY 121 ACAGGAGTAATCTTTAAACAGGCTCTAACGGTCAATATCAATGAAACACACTATTCCTGTA 180
 Db 2444 ACAGGAGTAATCTTTAAACAGGCTCTAACGGTCAATATCAATGAAACACACTATTCCTGTA 2503
 QY 181 AGTAATGAAATGCTGTTGCTAAATTTACCTCTCCAGGAGCTATTTTATGAAAGCTAGT 240
 Db 2504 AGTAATGAAATGCTGTTGCTAAATTTACCTCTCCAGGAGCTATTTTATGAAAGCTAGT 2563
 QY 241 TTTTATTAATATATTTTCAAGTTTGGCAGATATTAGTATAATGTTAAAGAAATGAA 300
 Db 2564 TTTTATTAATATATTTTCAAGTTTGGCAGATATTAGTATAATGTTAAAGAAATGAA 2623
 QY 301 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 360
 Db 2624 CAACACCAAGTGTGTTTAAACAGGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 2683
 QY 361 GACCAATATCTCTGCTACAGAGATGTAATCAAGAAATCTTTGATTTTAAACAA 420
 Db 2684 GACCAATATCTCTGCTACAGAGATGTAATCAAGAAATCTTTGATTTTAAACAA 2743
 QY 421 TTATTGAAGTCTATTTTCAAGTGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 480
 Db 2744 TTATTGAAGTCTATTTTCAAGTGTGTAATACAGAGATTACCTTTAAAGGAAAGATGTT 2803
 QY 481 ATTTTAAACAGGAGTTCATATTTGATTAAGTAATCAATGAAAGATTTTAAAGCAGTAGGACT 540

Db 2804 ATTTTACAGGAGTTCATATTGTTATTAAGCAATCATATAAGATTTTAAAGCAGTAGCGACT 2863
 Qy 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGAC-AATACTTCAGCAGATTTCAT 599
 Db 2864 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGACAAATATCTTCAGCAGATTTCGA 2923
 Qy 600 GGTAGTCTTCAAGTAATCTTTGAGAGAAATTTTCAGCAGATTTCAGCAGATTTCGA 659
 Db 2924 TGTGGTTATTCACAGTAATCTTTTCAGAGAAATTTTCAGCAGATTTCAGCAGATTTCGA 2983
 Qy 660 GACCGTTCAGGATTTTCTACCAAGCCAAATCTGTTTCAGAGTGAACATTTCTTT 719
 Db 2984 GACCGTTCAGGATTTTCTACCAAGCCAAATCTGTTTCAGAGTGAACATTTCTTT 3043
 Qy 720 TTATACAGCGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAATT 779
 Db 3044 TTATACAGCGCTCTTAGAAGGAAATTTATCCCGATACAGATCGTTTATTAATGACAGAATT 3103
 Qy 780 TGAGCAGGAGTGTGTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 839
 Db 3104 TGAGCAGGAGTGTGTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 3163
 Qy 840 GATTCTTAATGCTACTCAAAATGCTTAAAGCTT 876
 Db 3164 GATTCTTAATGCTACTCAAAATGCTTAAAGCTT 3200

RESULT 7
 ACA47964
 ID ACA47964 standard; DNA; 1134 BP.
 XX AC ACA47964;
 XX 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #29621.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX Streptococcus mutans.
 XX WO200277183-A2.
 PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU44094.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 35834; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1134 BP; 380 A; 194 C; 179 G; 381 T; 0 U; 0 Other;
 Query Match 60.9%; Score 690.8; DB 8; Length 1134;
 Best Local Similarity 75.6%; Pred. No. 7.9e-138;
 Matches 857; Conservative 0; Mismatches 277; Indels 0; Gaps 0;
 Qy 1 ATGATTCAATTTTCAATTAATCGCACATTATTATTCTTTTAAATACACTAAAGCT 60
 Db 1 ATGATAAATTTTCAATTAATTAAGTTTTTCTTACAGCCTTAATATGCTACCAAGCGA 60
 Qy 61 GCTATTAGCATAAAATGCCATTCTTATTTCTTTCATCAATAAAATTTGAAGTCACTTCT 120
 Db 61 GCTATTAGTCTTAAATGCTATTCTTCTTCTAGTTTAAATTTGAAGTGAATTTCT 120
 Qy 121 ACAGAGTAACTTTAAACAGGGTCTAACGGTCAATATCAATTAATGAAAACACTATTCTGTGA 180
 Db 121 CAGTCCATTACTTTAAACAGGGCTCTAATGAGCAAAATTTCTATTGAAAATACGATTTTCAGCT 180
 Qy 181 AGTAATGAATGCTGTTTCTTAATACCTCTCCAGGAGCTATTATTTATTAGAAGCTAGT 240
 Db 181 GAAGAAGAAATGCTGGACTATTGTTACTTCTTCAGGAGCTATCTTACTTGAAGCTAAT 240
 Qy 241 TTTTATTATAATATTATTCAAGTTTGGCAGATATTAGTATAAATGTTTAAAGAAATTCGA 300
 Db 241 TTTTATTATAATATTATTCAAGCTCTACCTGATATCACTTTAGATTTTGAAGAAATTCGA 300
 Qy 301 CAACACCAAGTTGTTTAAACCAAGTGGTAAATCAGAGATTACCTTAAAGGAAAAGATGTT 360
 Db 301 CAACATCAAGTTGTTTAAATAGTGGAAAATCAGAAATTTACACTTTAAAGGTAAGATGTG 360
 Qy 361 GACAGTATCTCGTCTACAGAGAGTATCAACAGAAATCCCTTGAATTTTAAAGCAAAA 420
 Db 361 GAACAAATATCTCGTTGAGGAAAGTTGGAACAAATATCTTTTAAATTTTGAAGCAAAA 420
 Qy 421 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCAGCAGCTTTTCAAGAAAGTCTCTCT 480
 Db 421 TTGTTAAAAACAATATTCTTGAACAGCTTTTTCGCTAGTACCCAGAAAGTCTGACCA 480
 Qy 481 ATTTTAAACAGAGTTTCATATTGTTAATAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
 Db 481 ATTTTAAACAGAGTTTCATTTGGTATTGACAAATCATAAAGAGTTTAAAGCTGTTCTGCTACA 540
 Qy 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGAACAATCTTTCAGCAGATTTCATG 600
 Db 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGAACAATCTTTCAGCAGATTTCATG 600

Qy 601 GTAGTCTTCCAGTAATCTTTGAGAGAAATTTTCAGCAGTATTTCAGAGATGATATTGAG 660
 Db 601 GTCGTCAATCCAGCCGTTCTTTACGTAATTTCCAGCTGATTTACAGATGATATTGAA 660
 Qy 661 ACCGTTGAGTATTTTCTCACCAGCAAACTTTGTCAGAGTGAAGTGAACATTTCTTTT 720
 Db 661 TCTGTTGAGTATTTTCTCAGAGTCAAACTTTGTCAGAGTGAAGTGAATATATAGTCTT 720
 Qy 721 TATACAGCTCTTTAGAGGAATATTCGATACAGACCGTTTATTAATGACAGATTT 780
 Db 721 TATACAGCTCTTTAGAGGAATATTCGATACAGACCGTTTATTAATGACAGATTT 780
 Qy 781 GAGAGGAGTGTGTTTCAATACCAATCCCTTCGCGACGCTATGGAAGCTGCTTCTTG 840
 Db 781 GAAACAGAGTGTGTTTCAATACCAATCCCTTCGCGACGCTATGGAAGCTGCTTCTTG 840
 Qy 841 ATTCTTAATGCTACTCAATGCTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 900
 Db 841 ATTCTTAATGCTACTCAATGCTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 900
 Qy 901 GCTCATGTTAACTCACTGAGTGTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 960
 Db 901 GCACATGTTAACTCACTGAGTGTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 960
 Qy 961 TCTGTTAGTATTTAACTCACTGAGTGTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1020
 Db 961 TCAGGAATGATTTAACTCACTGAGTGTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1020
 Qy 1021 ATTAAAGTGAACAGTAAATTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1080
 Db 1021 CTTAAAGTGAACAGTAAATTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1080
 Qy 1081 CCAGGCGATGAGGAAGAAAGTATTTTCACTTAACTTAACTTAACTTAACTTAACTTAA 1134
 Db 1081 CCAGGCGATGAGGAAGAAAGTATTTTCACTTAACTTAACTTAACTTAACTTAACTTAA 1134

RESULT 8
 ABN68600
 ID ABN68600 standard; DNA; 1134 BP.
 AC ABN68600;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Streptococcus polynucleotide SEQ ID NO 5113.
 XX
 KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR P-P8DB; ABP27969.
 XX

PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 7; Page 3674-3675; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 1134 BP; 381 A; 196 C; 186 G; 371 T; 0 U; 0 Other;
 Query Match 56.5%; Score 641.2; DB 6; Length 1134;
 Best Local Similarity 72.8%; Pred. No. 3e-127;
 Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
 Qy 1 ATGATTCAAATTTTCAATTAATCGCAGTATTATTTATTCATGCTTTTAAATACAACTTAAACGT 60
 Db 1 ATGATTCATTTTCAATAAATAAATAATTTCTTTCGATGCTTAAACGTAACCAACGA 60
 Qy 61 GCTATTAGACTAAAAATGCCATTCCTTATTTTCATCAATAAAAAATGAAGTCACATCTCT 120
 Db 61 GCTATTAGTCAAAAAATGCCATTCCTTATTTTCATCAATAAAAAATGAAGTCACATCTCT 120
 Qy 121 ACAGGAGTAACTTTAAACAGGCTTAAAGGTCATTAATCAATTCATTAACAACTATTCCTGTA 180
 Db 121 GATGCTATCATTTTAAACGGGTCAAATGGACAAATTTCAATGAAATATCTATTCCTGCT 180
 Qy 181 AGTAATGAAATGCTGTTTGTGTAATTAATCACTCTCCAGGAGCTATTTTATTAAGAGCTAGT 240
 Db 181 TCAAAATGAAATGCTGTTTGTGTAATTAATCACTCTCCAGGAGCTATTTTATTAAGAGCTAGT 240
 Qy 241 TTTTATTAATTAATTTTCAAGTTTCCAGATTTAGTATTAATTTAAAGAAATTTGAA 300
 Db 241 TTTTATTAATTAATTTTCAAGTTTCCAGATTTAGTATTAATTTAAAGAAATTTGAA 300
 Qy 301 CAACACCAAGTGTGTTTAAACAGTGTGTAATTAATCAAGATTAACCTTTAAAGGAAAGATGTT 360
 Db 301 CAACATCAATGTTTCTTACTAGTGAATTAATCAAGATTAACCTTTAAAGGAAAGATGTT 360
 Qy 361 GACCATGATCTGCTTCAAGAGTATCAACAGAAATTCCTTTGATTTTAAACAAACAAA 420
 Db 361 GATCAATACCTGCTTCAAGAGTATCAACAGAAATTCCTTTGATTTTAAACAAACAAA 420
 Qy 421 TTAATGAAGTCTATTTTCAAGTGTGTAATTAATCAAGATTAACCTTTAAAGGAAAGATGTT 480
 Db 421 CTGTTTAAATCAATTAATTAATGAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 480
 Qy 481 ATTTTAAACAGGAGTTCATATTTGATTAAGTAAATCAATAAGATTTTAAACAGGAGTGTG 540
 Db 481 ATCTTAAACAGGAGTTCATATTTGATTAAGTAAATCAATAAGATTTTAAACAGGAGTGTG 540
 Qy 541 GACTCTCATGATGAGCAAGCTTTTAACTATCACTTTGAGCAATTAATTTTAAAGATTTGATG 600
 Db 541 GATTCACCGTATGAGTCAACGCTTTTCCAAATTAGAGAAATCGGCTAATAATTTTAT 600
 Qy 601 GTAGTTCCTTCCAAAGTAAATCTTTTGAGAGAAATTTTTCAGAGATTTTACAGATGATTTAG 660
 Db 601 TTGGTTGTTCCAAAGTAAATCTTTTGAGAGAAATTTTTCAGAGATTTTACAGATGATTTAG 660

XX AC	ABX08064;	QY	1	ATGATTCAATTTTCAATTAATGCGCATTATTTATTCATGCTTTAAATACAACTAAACGT	60
XX DT	27-OCT-2003 (revised)	Db	1	ATGATTCAATTTTCAATTAATGCGCATTATTTATTCATGCTTTAAATACAACTAAACGT	60
XX DT	11-FEB-2003 (first entry)				
XX DE	S. pneumoniae type 4 strain coding region #2352.	QY	61	GCTATTAGCAGTAAAGTCCATTCCTATTCCTTTTCATCAATAAATTTGAAGTCACTTCT	120
XX KW	Gene; ds; bacterial meningitis; pneumonia; sepsis; otitis media;	Db	61	GCTATTAGTCTTAAAGTCCATTCCTATTCCTTTTCATCAAGTAAATTTGAAGTCACTTCT	120
XX KW	ear infection; antinflamatory; antibacterial; immunostimulant;	QY	121	ACAGGAGTAACCTTTAAACAGGCTTAACGCTCAATATCAATTTGAAACACATTTCTCTGA	180
XX KW	auditory; respiratory; gene therapy; vaccine.	Db	121	GAAGGATTAACCTTTAAACAGGCTTAACGCTCAATATCAATTTGAAACACATTTCTCTGA	180
XX OS	Streptococcus pneumoniae; type 4 strain.	QY	181	AGTAATGAAATGCTGGTTTCTAATACCTCTCCAGAGCTATTTTATTAAGAGCTAGT	240
XX PN	WO200277021-A2.	Db	181	AAAAATGAAGATGCTGGTTTCTAATTAATTAATTAATTAATTAATTAATTAATTAAT	240
XX PD	03-OCT-2002.	QY	241	TTTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	300
XX PF	27-MAR-2002; 2002WO-1B002163.	Db	241	TTCTTTATCAATGATGATCTAGTTTACCTGATGTAATCTCTTGAATTTAAAGAAATTTGA	300
XX PR	27-MAR-2001; 2001GB-00007658.	QY	301	CAACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGATGTT	360
XX PA	(CHIR-) CHIRON SPA.	Db	301	CAAAATCAAAATTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGATG	360
XX PA	(GENO-) INST GENOMIC RES.	QY	361	GACCAATATCCTGCTTACAGAGATGATCAAGAAATCCTTTGATTTTAAAGCAAAA	420
XX PI	Masignani V, Tettelin H, Fraser C;	Db	361	GAACAATATCCTGCTTACAGAGATGATCAAGAAATCCTTTGATTTTAAAGCAAAA	420
XX DR	WPI; 2003-040579/03.	QY	421	TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGAGCCAGTTTACAGAGATGCTGCT	480
XX DR	P-PSDB; ABU02772.	Db	421	TTACTCAAGAAATTTAATGAAACAGCTTTTGCTGCAAGTACACAGAGATGCTGCTG	480
XX PT	New proteins and nucleic acid molecules from Streptococcus pneumoniae,	QY	481	ATTTTAAACAGAGTTTCAATTTGATTAAGTAAATCATAAAGATTTTAAAGCAGTAGGACT	540
XX PT	useful as medicaments for treating or preventing a disease or infection	Db	481	ATTTTAAACAGAGTTTCAATTTGATTAAGTAAATCATAAAGATTTTAAAGCAGTAGGACT	540
XX PT	due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or	QY	541	GACTCTCATGATGAGCCAAAGCTTTTAAATCACTTTGAGCAATTTTACAGAGATTTGATG	600
XX PS	Claim 6; SEQ ID NO 4703; 56pp; English.	Db	541	GACTCTCATGATGAGCCAAAGCTTTTAAATCACTTTGAGCAATTTTACAGAGATTTGATG	600
XX CC	The invention relates to a protein comprising or having at least 50%	QY	601	GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGAGATTTGAG	660
XX CC	identity to any of the 2469 amino acid sequences, identified in the	Db	601	GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGAGATTTGAG	660
XX CC	specification (available on a computer readable format), or its fragment,	QY	661	ACCGTTGAGGATTTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAG	720
XX CC	expressed from 2469 of 2489 identified DNA coding regions from the	Db	661	ACCGTTGAGGATTTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAG	720
XX CC	Streptococcus pneumoniae type 4 strain genomic sequence appearing as	QY	721	TATACACGCTCTTAGAGGAAATTTATCCGATACAGACGCTTTTAAATGACAGATTTT	780
XX CC	AB556454. Also included are an antibody which binds one of the proteins,	Db	721	TATACACGCTCTTAGAGGAAATTTATCCGATACAGACGCTTTTAAATGACAGATTTT	780
XX CC	treating a patient by administering the protein, DNA or antibody (in a	QY	781	GAGAGGAGGTTGTTTTCATATCCCAATCCCTTCCGACGCTATGGAACGCTGCTTCTTG	840
XX CC	composition), a kit comprising first and second primers, which are the	Db	781	GAGAGGAGGTTGTTTTCATATCCCAATCCCTTCCGACGCTATGGAACGCTGCTTCTTG	840
XX CC	nucleic acid cited above or fragments between nucleotides 8-100 of a	QY	841	ATTTCTAATGCTACTCAAAATGCTACTGTTAGCTTGAATTTACTCAAAATCATATTTCA	900
XX CC	sequence not defined in the specification, for amplifying a target	Db	841	ATTTCTAATGCTACTCAAAATGCTACTGTTAGCTTGAATTTACTCAAAATCATATTTCA	900
XX CC	the first primer is substantially complementary to the target sequence	QY	901	GCTCATGTTTAACTCACCTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTTGTTAGT	960
XX CC	and the second primer is substantially complementary to the target sequence	Db	901	GCTCATGTTTAACTCACCTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTTGTTAGT	960
XX CC	of the target sequence, and where the parts of the primers having	QY	961	TCGTGTTAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1020
XX CC	substantial complementarity define the termini of the target sequence to	Db	961	TCGTGTTAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1020
XX CC	be amplified, assay comprising contacting a test compound with the	QY	1021	ATTTAAAGTGAACAGTAAATTTTCTTTTATCACCAGTTTCGACCTACCTACCTAAC	1080
XX CC	protein, and determining whether the test compound binds to the protein	Db	1021	ATTTAAAGTGAACAGTAAATTTTCTTTTATCACCAGTTTCGACCTACCTACCTAAC	1080
XX CC	and a Streptococcus pneumoniae bacterium, where one or more genes	QY			
XX CC	encoding the proteins has been rendered inactive. The proteins, nucleic	Db			
XX CC	acid molecules, antibody and compositions are useful as medicaments for	QY			
XX CC	treating or preventing a disease or infection due to streptococcus	Db			
XX CC	bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis	QY			
XX CC	media or ear infection. They are also useful in developing vaccines,	Db			
XX CC	diagnostics and antibiotics. The methods are useful for identifying	QY			
XX CC	immunodominant proteins. The present sequence is one of the 2489	Db			
XX CC	identified coding region from the genomic sequence. Note: The sequence	QY			
XX CC	data for this patent did not form part of the printed specification, but	Db			
XX CC	was obtained in electronic format directly from WIPO at	QY			
XX CC	ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to	Db			
XX CC	standardise OS field)	QY			
XX SQ	Sequence 1134 BP; 384 A; 208 C; 182 G; 360 T; 0 U; 0 Other;	Db			

Query Match 53.6%; Score 608.2; DB 10; Length 1134;
 Best Local Similarity 71.1%; Pred. No. 3.3e-120;
 Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

```
QY 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATACACGAGTACGACAAA 1133
DB 1081 CCAGCAGATCTGACGAGAGCTTCATCGAGCTCAATACACGAGTTCTGACAAA 1133

RESULT 11
AAV52153
ID AAV52153 standard; DNA; 21338 BP.
XX
AC AAV52153;
DT 23-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae genome fragment SEQ ID NO:20.
XX
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS Streptococcus pneumoniae.
XX
FN WO9818931-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US019588.
XX
PR 31-OCT-1996; 96US-0029960P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Kunsch CA, Choi GH, Dillon PJ, Rosen CA, Barash SC, Fannon M;
PI Dougherty BA;
XX
DR WPI; 1998-272225/24.
XX
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus pneumoniae.
XX
PS Claim 1; Page 257-269; 1409pp; English.
XX
CC The present invention describes a computer readable medium which has the
CC nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (AAV52134 to AAV52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S. pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1 to
CC 391, identifying members of the library which contain sequences that
CC hybridise to the target sequence and isolating the nucleic acid molecules
CC from the members; or (b) isolating mRNA, DNA or cDNA produced from an
CC organism, amplifying nucleic acid molecules whose nucleotide sequence is
CC homologous to amplification primers derived from the fragment of the S.
CC pneumoniae genome to prime the amplification and isolating the amplified
CC sequences. The computer readable medium can be used in a computer-based
CC system for identifying fragments of the S. pneumoniae genome of
CC commercial importance, or expression modulating fragments of the S.
CC pneumoniae genome. Products from the present invention can be used in
CC diagnosis kits and assays, and pharmaceutical compositions and vaccines
CC for S. pneumoniae
XX
SQ Sequence 21338 BP; 7339 A; 3446 C; 4112 G; 6441 T; 0 U; 0 Other;
Query Match 53.6%; Score 608.2; DB 2; Length 21338;
Best Local Similarity 71.1%; Pred. No. 5.5e-120;
Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;
QY 1 ATGATTTCAATTTTCAATTAATCGCACATTTATTTATTTATTTTAAATACAACTAAACGTT 60
DB 7076 ATGATTTCAATTTTCAATTAATTAATAAATTTATTTCTACAAAGCATTAATACTACTAAGAGA 7135
```

```
QY 61 GCTATTAGCACTAAAAATGCAATTCCTATTCTTTTCATCAATAAAAAATGAAGTCACTTCT 120
DB 7136 GCTATTAGTTCTAAAAATGCCATTCCTATTATTATCAACAGTAAAAAATGAGGTGACCAAT 7195
QY 121 ACAGGAGTAACTTTAAACAGGCTCTAAACGGTCAATATCAATTAATGAAAAACACTATTCTCTGA 180
DB 7196 GAAGGTATTACTTTAAATGGTTCAAAATGGTCAAAATTTCAATTTGAAAAATTTTATTCTCAA 7255
QY 181 AGTAATGAAAAATGCTGGTTTCTCTAATTAACCTCTCCAGGAGCTATTTTTATTAGAACCTAGT 240
DB 7256 AAAAATGAAGATGCTGGTTTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7315
QY 241 TTTTATTATATATATTTTAAAGTTTGCCAGATATTAGTATATAATGTTTAAAGAAATGAA 300
DB 7316 TTCTTTTCAATGTAGTATCTAGTTTACCTGATGTAACCTCTTGATTTTAAAGAAATGAA 7375
QY 301 CAACACCAAGTTGTTTAAACCAAGTGGTAAATCAGAGATTACCTTAAAGAGGAAAGATGTT 360
DB 7376 CAAATCAAAATGTTTAAACCAAGTGGCAATCAGAAATTTACCTTAAAGAGGAAAGATAGC 7435
QY 361 GACCAATATCTCGTCTACAAAGAGTATCAACAGAAAAATTCCTTTGATTTTAAAAACAAA 420
DB 7436 GAACAATATCCAGAAATCCAGAAATTTTACAGACGACTCTCTTTAATTAATTAATTAATTAAT 7495
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCTCT 480
DB 7496 TTACTCAAGAAATTTAATTAATGAAACAGCTTTTGTCTGCAAGTACACAAGAGAGTCTGCTCG 7555
QY 481 ATTTTAAACAGGATTCATATTGATTAAGTATCATAAAGATTTTAAAGCAGTACGAGT 540
DB 7556 ATTTTAAACAGGATTCATATTGATTAAGTATCATAAAGATTTTAAAGCAGTACGAGT 7615
QY 541 GACTCTCATCTGATGAGCCAAACGTTTAAATCACTTTGGACAATATCTTACAGCAGATTTGATG 600
DB 7616 GACTCTCATCTGATGAGCCAAACGTTTAAATCACTTTGGACAATATCTTACAGCAGATTTGATG 7675
QY 601 GTAGTTCTTCAAGTAAATCTTTGAGAGAAATTTTACAGCAGTATTTTACAGATGATATTGAG 660
DB 7676 GTCTGTAATTCCTAGCCGTTCTCTACGCGAATTTTACGCGGATTTTACAGATGATATTGAG 7735
QY 661 ACCGTGAGGATTTTCTCTACCAAGCCAAATCTTTTCAAGAGTGAACACATTTCTTTT 720
DB 7736 ACTGTAGAGATTTTCTTTTGCCAAATACCAAAATCTCTTTTGAAGGAAATTAATTAAGCTTC 7795
QY 721 TATACAGCCCTCTTTAGAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
DB 7796 TATCTGCTCTCTAGAGGAAATCTATCTCTGATACAGATCGCTTGATTTCCACAGACTTT 7855
QY 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCCGACCGCTATGGAACGCTGCTCTTTG 840
DB 7856 AACACTACTATTACTTTTAAATGCGTAACTTACGCCAGTCAATGGAGCGTGCCTGCTTT 7915
QY 841 ATTTCTAATGCTACTCAAAATGCTAGTAAAGTGAAGTAAACGAGGATTTAGATATTGTTAGTCA 900
DB 7916 TTATCAAGTGGAGTCAAAATGGTACTGTGAAACTTTGAAATTAAGAGTGGGGTGTGTAGC 7975
QY 901 GCTCATGTTAACTCACCTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCA 960
DB 7976 GCCCATGTTCACTCTCCAGAGTTGGTAAAGTAAACGAGGAAATCGATCTGATCAGTT 8035
QY 961 TCTGGTGTGATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 1020
DB 8036 ACTGGTGAAGATTTGACCATTAGTTTCAACCCAACTTACTTGTGATTTCTTCTTAAAGCT 8095
QY 1021 ATTTAAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTCGACCAATTCACCCCTAACA 1080
DB 8096 TTAATACGGAAGAGGAGTACTATTAGCTTTTATCTCAGCTGTTCGTCCTTACTCTTGTG 8155
QY 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1133
DB 8156 CCAGCAGATCTGACGAGAGACTTTCATGAGCTCATTACACGAGTTCGTTACAAA 8208
```


Qy	601	GTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTCAG	660
Db	56827	GTGCTAAATCTCTAGCGGTTCTCTACGCGAAATTTTTCAGCGGATTTTACAGATGATATTCGAA	56886
Qy	661	ACCGTTGAGGTATTTTCTCACCAAGCCAAATCTTGTTCAGAAAGTGAACAAATTTCTTTT	720
Db	56887	ACTGTAGAGATTTCTTTCGCAATTAACCAATCTCTTTTAGAAGCGAAATATTTAGCTTTC	56946
Qy	721	TATACACGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAATTT	780
Db	56947	TATACTCGTCTCTAGAAGGAAATATCTGTATACAGATCGCTTGATTTCCACAGACTTT	57006
Qy	781	GAGACGGAGGTCTTTTTCATACCCCAATCCCTTCGCCACGCTATCGAAGCGCTTCTTCTTG	840
Db	57007	AACACTACTATTTACTTTTAAATGTGGTAAACTTTACGCCAGTCAATGGAGCGTGCCTCTT	57066
Qy	841	ATTTCTAAATGCTACTCAAAATGGTACTGTAAAGCTTTGAGATTACTCAAAAATCATATTTCA	900
Db	57067	TTATCAAGTCGACTCAAAATGGTACTGTGAACTTGAATTTAAGATGGGTTGTTTACG	57126
Qy	901	GCTCATGTTAACTCACCTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTTGTAGTCAG	960
Db	57127	GCCCATGTTTCACTCTCCAGAAGTTGGTAAAGTAAACGAGGAAATCGATACTGATCAGTT	57186
Qy	961	TCGTGTAGTGATTAACTACTATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTTAAAGCT	1020
Db	57187	ACTGGTGAAGATTGGACATTAGTTTCAACCCAACTTACTTGATTGATTTCTTTAAGCT	57246
Qy	1021	ATTTAAAGTGAACAGTAAAAATTCATTTCTTTATCACCGATTCGACCATTCACCTTAACA	1080
Db	57247	TTAAATAGCGAAAGGTGACTATTAGCTTTATCTCAGCTGTTCGTCCTTTACTCTTGTG	57306
Qy	1081	CCAGGCGATGAGCAAGAGTTTTTATCCAAATTAATACACCATGATCAGCAACAAA	1133
Db	57307	CCAGCAGATCTGACAGACTTTCAGCAGCTCATACACAGTTTCGTACAAA	57359
RESULT 13			
ID	AAS55711 standard; DNA; 1137 BP.		
AC	AAS55711;		
DT	13-FEB-2002 (first entry)		
DE	Streptococcus pneumoniae DNA for cellular proliferation protein #282.		
KW	Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;		
XX	antibacterial; drug design.		
OS	Streptococcus pneumoniae.		
PN	WO200170955-A2.		
PD	27-SEP-2001.		
PF	21-MAR-2001; 2001WO-US009180.		
PR	21-MAR-2000; 2000US-0191078P.		
PR	23-MAY-2000; 2000US-0206848P.		
PR	26-MAY-2000; 2000US-0207727P.		
PR	23-OCT-2000; 2000US-0242578P.		
PR	27-NOV-2000; 2000US-0253625P.		
PR	22-DEC-2000; 2000US-0257931P.		
PR	16-FEB-2001; 2001US-0269308P.		
PA	(ELIT-) ELITRA PHARM INC.		
PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;		
FI	Yamamoto RT, Xu HH;		
DR	WPI; 2001-611495/70.		
DR	P-PSDB; AAU37852.		

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
PS
XX Claim 27; SEQ ID NO 9348; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1137 BP; 384 A; 210 C; 182 G; 361 T; 0 U; 0 Other;
Query Match 53.2%; Score 603.4; DB 4; Length 1137;
Best Local Similarity 70.8%; Pred. No. 3.5e-119;
Matches 802; Conservative 0; Mismatches 331; Indels 0; Gaps 0;
QY 1 ATGATTCATTTTCAATTAATGCGACATATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 1 ATGATTCATTTTCAATTAATTAATAAATTTATTTCTACAAGCATTAATTAATTAAGAGA 60
QY 61 GCTATTGACACATAAATGCGATTCCTATTTCTTCATCAATAAATAATGAGTCACTTCT 120
DB 61 GCTATTGATTTCTAAATAATGCGATTCCTATTTATCACTGTCATAAATGAGTCACTTCT 120
QY 121 ACAGGATTAATTTTAAACAGGGTCTAACCGTCAATATCAATTTGAAACACATTTCTCTGA 180
DB 121 GAAGGTGTTACTTTAATTTGTTTCAATGTCATATTTCAATTTGAAATTTATTTCTCAA 180
QY 181 AGTAATGAATGCTGGTTGCTTAATACCTTCCAGGAGCTATTTTATTAAGAGCTAGT 240
DB 181 AAAAATGAAGATGCTGGTTGCTTAATTTACTTTTAGGTTGATCCTTTCTTGAAGCTTCT 240
QY 241 TTTTATTAATATTTATTTCAAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATGAA 300
DB 241 TTTTATTAATATTTATTTCAAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATGAA 300
QY 301 CAACACCAAGTTGTTTAAACAGGTTGTAATCAGAGATTACCTTAAAGGAAAGATGTT 360
DB 301 CAAATCAATGTTTAAACAGGTTGTAATCAGAGATTACCTTAAAGGAAAGATGTT 360
QY 361 GACAGATATCTCGTCTCAAGAAGTATCAACAGAGAAATCTTTGATTTTAAACAAAA 420
DB 361 GAAATCAATATCCAGGAATCCAGAGAAATTTTCAAGAGCTCTTTTAAATTTTAAACAAAA 420
QY 421 TTATTTGAGTCTATTTATTTGCTGAACAGCTTTTGCAGGCTTTTCAAGAGAGTCTGCT 480
DB 421 TTATTTGAGTCTATTTATTTGCTGAACAGCTTTTGCAGGCTTTTCAAGAGAGTCTGCT 480
QY 481 ATTTTAAACAGAGTTTCATATTTTAAAGTATATCAATAAAGATTTTAAAGCAGTACGACT 540
DB 481 ATTTTAAACAGAGTTTCATATTTTAAAGTATATCAATAAAGATTTTAAAGCAGTACGACT 540
QY 541 GACTCTCATCGTATGAGGCAACGTTTATTAATCTTTTGGACAAATCTTCAGCAGATTTGATG 600
DB 541 GACTCTCATCGCTAAGCCAGAGAAAAATTAATCTTTGAAAAAATAAGTATGATTTTATGAT 600
QY 601 GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATTTAGAG 660
DB 601 GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATTTAGAG 660

DB 601 GTCGTAATTCCTAGCCGTTCTCTAGCGAATTTTTCAGCGGTATTTTACAGATGATATCGAA 660
QY 661 ACCGTTGAGGTATTTTCTCACCAGGCCAAATCTTTGTTGTTGTTGTTGTTGTTGTTGTT 720
DB 661 ACTGTAGAGATTTTCTTTTGGCCAAATTAACCAATCTCTTTTAGAAGCGAAATATTAGCTTC 720
QY 721 TATACACGCTCTTTAGAGAGAAATTTATCCCATATACAGACCGTTTATTTATGACAGAAATTT 780
DB 721 TATATCTGCTCTCTAGAGAGAAATTTATCTCTGATACAGATCGCTTTGATTTCCAACAGACTTT 780
QY 781 GAGACGGAGGTTGTTTCTCAATACCAATCTCTGCGCACGCTATTCGAAACGCTCTTCTTCTG 840
DB 781 AACACTATATTTACTTTTAAATGTTGTAACCTTACCCAGTCAATGAGGCGGCTCTTCTT 840
QY 841 ATTTCTATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTTCAAAATCATATTTCA 900
DB 841 TTATCAAGTCGACTCAAAATGGTACTGTTGAACTTTGAAATTAAGGATGGGTTCTGTAGC 900
QY 901 GCTCATGTTAACTCACTGAGGTTGTTAAGGTTAAGGTTAAGGTTAAGGTTAAGGTTAAGGTT 960
DB 901 GCTCATGTTAACTCACTGAGGTTGTTAAGGTTAAGGTTAAGGTTAAGGTTAAGGTTAAGGTT 960
QY 961 TCTGTAGTGAATTTAACTATCATGCTTCAATCCAACTTCACTTATTTAGTCTTTTAAAGCT 1020
DB 961 ACTGTGAGATTTGACCAATTTAGTTTCAACCACTTACTTGAATTTGATTTCTTTAAGCT 1020
QY 1021 ATTTAAAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCCAGCATTCCACCTTAA 1080
DB 1021 TTTAAATGCGAAAGGTTGACCATTTAGCTTTTATCTCAGCTGTTTCTGTTCTTCTGTTG 1080
QY 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATTTACACGATGTTACGAGTACGAGCAAA 1133
DB 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATTTACACGATGTTACGAGTACGAGCAAA 1133
RESULT 14
ACA49628
ID ACA49628 standard; DNA; 1137 BP.
XX
AC ACA49628;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #31285.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus pneumoniae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
(ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
WIPI; 2003-029926/02.
DR P-PSDB; ABU45758.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids, required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Sd

Claim 14; SEQ ID NO 37498; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/publ/published_pct_sequences

Sequence 1137 BP; 384 A; 210 C; 182 G; 361 T; 0 U; 0 Other;

Query Match	53.2%	Score 603.4;	DB 8;	Length 1137;
Best Local Similarity	70.8%	Pred. No. 3.5e-119;		
Matches 802; Conservative	0;	Mismatches 331;	Indels 0;	Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCACATATTATTATTCATGCTTTAAATACAACCTAAACGT 60

Dbb
1 ATGATTCATTTTCAATTATATAAAATTTATTTCTCAAGCATTTAAATATTACTAAGAGA 60

Qy 61 GCTATTAGCACTAAAAATGCCATTCCTATTCTTTTCATCAATAAAATTTGAAGTCACCTTCT 120

Db 61 GCTATTAGTTCTAAAAATGCCATTCTATTTATCAACTGTCAAAATTGACGTGACCAAC 120

QY
121 ACAGGAGTAACTTTAAACAGGGTTAAACGGTCAAATATCAATTGAAAAACACATATTCCTGTA 180

Db
121 GAAGGTGTTACTTTAAATTGGTTCAAATGGTCAAATTTCAAATGAAAATTTTATTCTCAA 180

QY
181 AGTAATGAAAATGCTGGTTTGCTAAATTACCTCTCCAGGACCTATTTTATTAGAACCTAGT 240

Db
181 AAAAATGAAGATGCTGGTTGCTTAATTACTCTTAGGTCGATCCTCTTGAAGCTTCT 240

241 TTTTATTAATATTATTTCAAGTTGCCAGATATTAGTATAAATGTTAAGAAATTGAA 300

Db
241 TTCTTTATCAATGTAGTATCTAGTTTACCTGATGTAACCTCTGATTTAAAGAAATTGAA 300

301 CAACACCAAGTTGTTTTAACCGAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360

Db
301 CAAAATCAAAATGTTTTAACCAAGTGGCAAAATCAGAAATTACCCATAAAGGAAGATAGC 360

361 GACCAGTATCCTCGTCTACAGGAAGTATCAACAGAAAATCCTTTTGATTTTTAAAAACAAA 420

Db
361 GAACAATATCCACGAATCCAGAAATTTTCAGCAAGCACTCCTTTAATACTTGAACAAAA 420

421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGAGCCAGTTTCAAGAAAGTCGTCT 480

Db 421 TTACTCAAGAAAATTATTAAATGAAACAGCCCTTTGCTGCAAGTACACAAGAGAGTGTGTCG 480

Qy	481	ATTTTAA	CAGAGTTCATATTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGTCGACT	540
Db	481	ATTTTAA	CAGGTGTCCACTTCGGTATTTGAGTCAACACAAAGAGTTTAAAAACAGTTTGCACA	540
Qy	541	GACTCTCAT	CTGTATGAGCAACGGTTTAACTCACTTTGGACAATACTTCACGACAGATTGTGAT	600
Db	541	GACTCTCAT	CGCTTAAGCCAGAAAATTAATCTTTGAAAAAATAGTGATGATTTTGAT	600
Qy	601	GTAGTCTT	CCAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATATTGAG	660
Db	601	GTCGTAA	ATCTCTAGCCGTTCTACGGGAAATTTTCAGCGGTATTTACAGATGATATCGAA	660
Qy	661	ACCGTTGAG	GTATTTTCTCACCAAGCCAAATCTGTTTCAGAAGTGAACACATTTCTTTT	720
Db	661	ACTGTAGAG	ATTTCTTTGCGAAATACCAATCTCTTTAGAACGGAATAATATTAGCTTC	720
Qy	721	TATACAG	CCCTCTTAGAAGGAAATTAATCCCGATACAGACCCGTTTATTAATCACAGAAATTT	780
Db	721	TATACTG	CTCTCTAGAAGGAAACTATCTCTGATACAGATCGCTTGATTTCCAACAGACTTT	780
Qy	781	GAGACGGAG	GTGTTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCCTTG	840
Db	781	AACACTACT	ATTAATGCTGTTAACTTACGCCAGTCAATGGAGCGTGCCTGCTT	840
Qy	841	ATTTCTA	ATGCTACTCAAAATGGTACTGTTAAAGCTTCGAGATTACTCAAAATCATATTCA	900
Db	841	TTATCAAG	TGCGACTCAAAATGGTACTGTGAACTTGAATTAAGGATGGGTTGTTAGC	900
Qy	901	GCTCATG	TTAACTCACCTGAGGTGGTAAAGTAAACGAGGATTTAGATATTTAGTTCAG	960
Db	901	GCCCATG	TTCACCTCCAGAAAGTTGGTAAAGTAAACGAAGAAATCGATCTGATCAGTT	960
Qy	961	TCCTGTAG	TGATTTAACTATACAGCTTCAATCCCACTTACCTTATGAGTCTTTTAAAGCT	1020
Db	961	ACTGCTGA	AGTTGACATTAGTTTCAACCCAACTTACTTGATTTGATTTCTTTAAAGCT	1020
Qy	1021	ATTTAAAG	TGAACAGTAAAAATTCATTTCTTATCACAGTTCGACCATTCACCTTAACA	1080
Db	1021	TTAAATAG	CAAAAGGTGACCAITAGCTTTATCTCAGCTGTTCCTCCATTACTCTTTG	1080
Qy	1081	CCAGCGCAT	GAGGAGAAAGTTTATCCCAATTAATACACAGTACGAA	1133
Db	1081	CCAGCAGAT	ACTGACGAGCTTCACTGACGCTCATTACACAGTTCGTCAAA	1133
RESULT 15				
ABA90521_00				
WP	Sequence split into 24 fragments			
WP	Fragment Name	Begin	End	Accession
WP	ABA90521_00	1	110000	ABA90521
WP	ABA90521_01	100001	210000	
WP	ABA90521_02	200001	310000	
WP	ABA90521_03	300001	410000	
WP	ABA90521_04	400001	510000	
WP	ABA90521_05	500001	610000	
WP	ABA90521_06	600001	710000	
WP	ABA90521_07	700001	810000	
WP	ABA90521_08	800001	910000	
WP	ABA90521_09	900001	1010000	
WP	ABA90521_10	1000001	1110000	
WP	ABA90521_11	1100001	1210000	
WP	ABA90521_12	1200001	1310000	
WP	ABA90521_13	1300001	1410000	
WP	ABA90521_14	1400001	1510000	
WP	ABA90521_15	1500001	1610000	
WP	ABA90521_16	1600001	1710000	
WP	ABA90521_17	1700001	1810000	
WP	ABA90521_18	1800001	1910000	
WP	ABA90521_19	1900001	2010000	
WP	ABA90521_20	2000001	2110000	
WP	ABA90521_21	2100001	2210000	
WP	ABA90521_22	2200001	2310000	
WP	ABA90521_23	2300001	2365589	

XX PF 21-MAR-2001; 2001WO-US009180.
 XX PR 21-MAR-2000; 2000US-0191078P.
 XX PR 23-MAY-2000; 2000US-0206848P.
 XX PR 26-MAY-2000; 2000US-0207727P.
 XX PR 23-OCT-2000; 2000US-0242578P.
 XX PR 27-NOV-2000; 2000US-0253625P.
 XX PR 22-DEC-2000; 2000US-0257931P.
 XX PR 16-FEB-2001; 2001US-0269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX PI Yamamoto RT, Xu HH;
 XX DR WPI; 2001-611495/70.
 XX DR P-PSDB; AAU35314.
 XX PT New polynucleotides for the identification and development of
 XX PT antibiotics, comprise sequences of antisense nucleic acids.
 XX PS Claim 27; SEQ ID NO 6810; 511pp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 XX CC prokaryotic cellular proliferation, their use in identifying the genes,
 XX CC their use in the discovery of novel antibiotics, the essential genes
 XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 XX CC useful for the identification of potential new targets for antibiotic
 XX CC development. The antisense nucleic acids can also be used to identify
 XX CC proteins used in proliferation, to express these proteins, and to obtain
 XX CC antibodies capable of binding to the expressed proteins. The proteins can
 XX CC be used to screen compounds in rational drug discovery programmes. The
 XX CC antisense nucleic acid sequence is also useful to screen for homologous
 XX CC nucleic acids which are required for cell proliferation in a wide variety
 XX CC of organisms. The present sequence encodes an essential prokaryotic
 XX CC cellular proliferation protein. Note: The sequence data for this patent
 XX CC did not form part of the printed specification, but was obtained in
 XX CC electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1131 BP; 385 A; 198 C; 193 G; 355 T; 0 U; 0 Other;
 Query Match 30.1%; Score 341.2; DB 4; Length 1131;
 Best Local Similarity 57.0%; Pred. No. 3.2e-63;
 Matches 644; Conservative 0; Mismatches 483; Indels 3; Gaps 1;
 QY 4 ATTCAATTTCATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGGTCT 63
 Db 1 ATGAAATTAACGGTCAACGAAGTCTTTTACGAAGATTACAACTGTTCAACGAGCA 60
 QY 64 ATTACACTAAATGCGATTCCTTATCTTCATCAATATAATTAAGTCACTTTTACA 123
 Db 61 ATTCTCTTAAACAAACGATTCCTTATCTTCATCAATATAATTAAGTCACTTTTACA 120
 QY 124 GGAGTAATCTTAACAGGCTCAACGGTCAATATCAATTAATTAAGTCACTTTTACA 183
 Db 121 GCTTATCACTTACTGAGGATTAACGGGATTAATTAATTAAGTCACTTTTACA 180
 QY 184 AATGAATGCTGGTTGCTAATTAACCTCTCCAGGAGCTATTTTATTAAGAGCTAGTTT 243
 Db 181 GATGAAAGCTCAATGACCATTTGAACGACAGGTAGCATCGTTTATCAATCTCGTTTC 240
 QY 244 TTTATTAATATTTTCAAGTTTCCAGATATTAGTATAATGTTAAAGAAATTAAGCAA 303
 Db 241 TTTGCGGAATATTCGTAATTAACGAGATATGTTCAATGGAAGTTCTAGATAAC 300
 QY 304 CACCAAGTTGTTTAAACAGTGGTAAATACAGATTTACCTTAAAGGAAAGATGTTGAC 363
 Db 301 AATCAAGTTCGAATTAATCTTACGAAAGAGCTGATTTTACGGTTAATGATTAAGTCTGAC 360

QY 364 CAGTATCTCTGCTCAAGAGAGTATCAACAGAAATCTCTTTGATTTTAAACAAATTA 423
 Db 361 AATTATCCGCACCTTACCAGTAATTCATCTCAAAACCAATGAATTTACCTGTTCAATTA 420
 QY 424 TTGAAGTCTATTATTGCTGAAACAGCTTTTGACGACGATTTTCAAGAAAGTCTCTATT 483
 Db 421 TTGACAAAATTTATTAGTGAACAGGTTTCTGTATCGATGACGAAAGTCTGCAATTT 480
 QY 484 TTAACAGGAGTTTCAATTTGATTAATTAAGTAAATCAATGAAGTTTAAAGCAGTACGACTGAC 543
 Db 481 TTAACCTGGGTTCACTTTATTTAGAAATCAAAA---TTACTTGGCGTTGCGACAGAT 537
 QY 544 TCTCATCTGATGAGCCAAAGTTTAACTTCACTTTGGCAATATCTTCAAGAGATTTGATGTA 603
 Db 538 TCACATCTGTTTAAAGTCAACGTTGATCCGACAGAAACAAAGCAGTAGAAGACTTTAAACATT 597
 QY 604 GTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTGAGACC 663
 Db 598 GTAATTCAGGAAAGTTTAACTGTAATTTCTCGTTCAATTAACCAATGAAGAAAGATG 657
 QY 664 GTTGAAGTATTTTCTCAACCAAGCCAAATCTTTGTCAGAAAGTGAACACATTTCTTTTAT 723
 Db 658 GTTGAATTCAGCATTTATGGAACCAAGTCTATTTTAAACAGAGAAACAATCTACTCTAT 717
 QY 724 ACAGCTCTTTAGAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTTGAG 783
 Db 718 TCTGTTTGTAGAGGAAATTTATCCTGATACCAACCGTTTAAATTCACACTAGCCATAAC 777
 QY 784 ACGAGGTTGTTTCAATCCCAATCTCTTCGACCGCTATGGAACGTCCTCTTCTGATT 843
 Db 778 ACGCAATTCGAATTTTATGACAGAAATGCTTTTCAGCAATCGAACGTCCTCTTACTT 837
 QY 844 TCTAATCTACTCAAAATGCTACTGTTAAGTAAAGTAAAGGATTTAGATATTGTTAGTCT 903
 Db 838 TCTCATGAAGAGCTGAACAAATTTGTCGCTTTCAATTTCCACAGATCTCTGTTGTTA 897
 QY 904 CATGTTAACTCACCTGAGGTTGCTGAAGTAAAGGATTAAGGATTTAGATATTGTTAGTCT 963
 Db 898 TATGGAATTCACCTGAAATTTGAAAGTGAAGAGCTTTAAACTATGAAGATGTTCT 957
 QY 964 GGTAGTATTAATCACTATCAGCTTCAATCCAACTTACCTTATTAAGTCTTTTAAAGCTATT 1023
 Db 958 GGTGAAGCGTTGATATTTCTTCAACCCAGATTTATATGAAGATCGTTTGGAGCGTTT 1017
 QY 1024 AAAAGTGAACAGCTAAATAATTCATTTCTTATCAGCTTCGACCATTCACCCCTAACCA 1083
 Db 1018 GGCATATGATTAATTAACCGTGAATTTCTTCTCAATTCGTCGCTTTACATTTGGAGCCA 1077
 QY 1084 GCGATGAGGAGAAAGTTTATCCAAATTAATTAACCAAGTACGAAACAAA 1133
 Db 1078 ACCGAAACAGAACTAGATTTTCAATTAATTAACCGGTTTCGTACAAA 1127

RESULT 17
 ACA32930
 ID ACA32930 standard; DNA; 1128 BP.
 XX AC
 AC ACA32930;
 XX XX
 DT 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #14587.
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene..
 XX OS Enterococcus faecalis.
 XX XX
 PN WO20027183-A2.
 XX PD 03-OCT-2002.
 XX XX
 PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR P-PSDB; ABU29060.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 20800; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1128 BP; 383 A; 196 C; 193 G; 356 T; 0 U; 0 Other;
 Query Match 29.8%; Score 338; DB 8; Length 1128;
 Best Local Similarity 56.8%; Pred. No. 1.5e-62;
 Matches 642; Conservative 0; Mismatches 485; Indels 3; Gaps 1;
 QY 4 ATTCAATTTCAATTAATCGCATATTTATTATTCATGCTTTTAAATACAACTAAACGTCT 63
 DB 1 ATGAAATTAACGGTCAAGGAGTGTCTTTTACAAAGAAATTAACAACATGTTCAACGAGCA 60
 QY 64 ATTAGCACTAAATAATGCCATCTCTTCTTCAATCAATAAAATTTGAAGTCACCTTCTACA 123
 DB 61 ATTTCTTTCAAAACAACGATTCATCTTAACAGGTGTAAATTTGTGTATCAGAAGAT 120
 QY 124 GGAGTAACCTTTAAACGGGTCTAACCGGTCAAATATCAATTTGAAACACACTATTCCTCTAAGT 183
 DB 121 GCGTTATCACTTACTCTGGGAGTAACCGGATATTTCAATTTGAAAGTTTTTAAAGTAAAGAT 180
 QY 184 AATGAAATGCTGGTTTGGTAAATACCTCTCCAGGAGCTATTTTATTAGAACCTAGTTT 243

Db 181 GATGAAAAGGCTCAAAATGACCAATTAACGACAGTAGCATCGTTTTTACAATCTCGTTTC 240
 QY 244 TTTATTAATATTTTCAAGTTTGGCAGATATTTAGTATATAATTTGTTAAAGAAATTTGAACAA 303
 Db 241 TTTGCGGAAATTTTCGTAATTTACCAAGAAGATATGTTTCACAATGGAAGTTCTAGATAAC 300
 QY 304 CACCAAGTTGTTTTAAACCAAGTGGTAAATACAGAGATTACCTTTAAAGGAAAGAGATTGTGAC 363
 Db 301 AATCAAGTTGCAATTTACTTTCAGGAAAGCTGATTTTACGGTTAATGGAATAGATCTGAT 360
 QY 364 CAGTATCTCGTCTACAGAAGATATCAACAGAAATCTTTTGATTTTAAACCAAAATTTA 423
 Db 361 AATATCCGCACTTACCAAGTAATGATACTCAAAACCAATGAAATTTACCTGTTCAATTTA 420
 QY 424 TTGAAGTCTATTTATGCTGAAACAGCTTTTGCAGCCAGTTTTCACAAAGAAGTCGTCTATT 483
 Db 421 TTGACAAAATTTATTTAGTGAACAGGTTTTGCTGTATCGATGCACGAAAGTCGTCCAATT 480
 QY 484 TTAACAGGAGTTTCAATTTGTTAATTAAGTAATCATTAAGATTTTAAAGCAGTAGCGACTGAC 543
 Db 481 TTAATCGGGTTCACTTTTATTTTAAATAATCAAAA---TTACTTTGCCGTTGCGCAGAT 537
 QY 544 TCTCATCTATGAGCAACGTTTAAATCACTTTGGACAATCTTTCAGCAGATTTGATGGTA 603
 Db 538 TCACATCGTTTAAGTCAACGTTGTATCCGACAGAACAGCAGTAGAAGACTTTAAACATT 597
 QY 604 GTTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTGAGACC 663
 Db 598 GTAAATCCAGGAAAAGTTTAACTTCTCGTTCAATTAACCAATGAAGAAGAAATG 657
 QY 664 GTTAGGATTTTCTTACCAAGCAATCTTTGTCAGAAAGTGAACACATTTCTTTTAT 723
 Db 658 GTTGAATTCAGCATTTATGGAATAATCAAGTGTCTTTTAAACAGAGAAACAAATGTAATCTAT 717
 QY 724 ACAGCCTCTTTAGAGGAAATTTATCCGATACAGACCGTTTATTATTAATGACAGAAATTTGAG 783
 Db 718 TCTGTTTGTAGAGGAAATTTATCCGATACCAACCGTTTAAATTCCAACTAGCCATAC 777
 QY 784 ACGGAGGTGTTTTCATACCAATCCCAATCCCTTCGCCACGCTATGGAAACGTCCTTCTTGATT 843
 Db 778 ACGCAATTTGAATTTTATGTACCAAGATTTGCTTTTCAGCAATCGAACGTCCTTCTTACTT 837
 QY 844 TCTAATGCTACTCAAAATGGTACTGTTTAAGCTTTGAGATTACTCAAAATCATATTTTCAGCT 903
 Db 838 TCTCATGAAGACGTAACAATATTTGTCGCTTTTCAATTTTCCAGATTTCTGTTGTTTA 897
 QY 904 CATGTTAACTCACCTGAGGTTTGGTAAGGTAAACAGGATTTAGATATTTGTTAGTCAGTCT 963
 Db 898 TATGAAATTCACCTGAAATTTGGAAGTGGAAAGCTTTAACTATGAAATGTTTCT 957
 QY 964 GGTAGTGAATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGCTTTTAAAGCTATT 1023
 Db 958 GGTGAAGGTTGGATATTTCTTTCAACCCAGATATATGAAAGATGCGTTGCGAGCGTTT 1017
 QY 1024 AAAAGTGAACAGTAAATAATTTCTTATACAGCTTCGACCACTTACCCCTAACACCA 1083
 Db 1018 GCGGATATGAATATTAACCGTGAATTTCTTCTCAATTCGTCGCTTTACATTTGAGGCCA 1077
 QY 1084 GCGGATAGGAGGAAAGTTTTTATCCAAATTAATTAACCCAGTACGAAACAAA 1133
 Db 1078 ACCGAAACAGAACTAGATTTTCACTCACTAATTAACCCGGTTGCTAGAAA 1127
 RESULT 18
 AAX13559
 ID AAX13559 standard; DNA; 8001 BP.
 XX
 AC AAX13559;
 XX
 DT 19-MAR-1999 (first entry)
 XX
 DE Enterococcus faecalis genome contig SEQ ID NO:622.
 XX

PF	04-MAY-1998;	98US-00070927.	
XX			
XX	04-MAY-1998;	98US-00070927.	
XX			
XX	(KUNS/) KUNSCH C A.		
PA	(DILL/) DILLON P J.		
PA	(BARA/) BARASH S.		
XX			
PI	Kunsch CA, Dillon PJ, Barash S;		
XX			
XX	WPI; 2002-750065/81.		
XX			
DR			
XX	Computer readable medium having recorded on it a Enterococcus faecalis		
PT	nucleotide sequence useful for detecting diseases related to Enterococcus		
PT	infections in animals.		
XX			
XX	Claim 1; Page; 119pp; English.		
XX			
CC	The present invention relates to a new computer readable medium with an		
CC	Enterococcus faecalis nucleotide sequence. The invention is useful to		
CC	diagnose the presence of E.faecalis in a sample or determining the		
CC	presence of a specific microbe in a sample. The invention is also useful		
CC	for modulating the growth or pathogenicity of E.faecalis, in a vaccine to		
CC	confer resistance to Enterococcal infection, for commercial, therapeutic		
CC	and industrial purposes, and for fermenting a particular sugar source or		
CC	to produce a particular metabolite. The invention is useful for detecting		
CC	diseases related to Enterococcus infections in animals, and for detecting		
CC	E.faecalis using biochip technology. The present nucleic acid sequence		
CC	represents an Enterococcus faecalis contig DNA sequence of the invention.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification but was obtained in electronic format directly from USPTO		
CC	at http://seqdata.uspto.gov		
XX			
SQ	Sequence 8001 BP; 2672 A; 1407 C; 1670 G; 2250 T; 0 U; 2 Other;		
	Query Match 29.2%; Score 330.6; DB 6; Length 8001;		
	Best Local Similarity 56.9%; Pred. No. 8.1e-61;		
	Matches 645; Conservative 0; Mismatches 485; Indels 4; Gaps 2;		
QY	1 ATGATTCAATTTTCAATTAAATCGCACATTATTATTTCATGCTTTAAATACAACCTAAACGT 60		
DB	598 ATTATGAAATTACGGTCNACGAGTGTCTTTTACAGAAATTACAACTGTTCAACGA 657		
QY	61 GCTATTAGCACTAAAAATGCGATTCTCTATTCTTTT-CATCAATAAAAATTGAAAGTCAC TTC 119		
DB	658 GCAATTTCTTCTRAAACACAGATTCCCAATCTTAACAGAGTGTAAAAATTGTTTATCAGA 717		
QY	120 TACAGGAGTAACTTTAAACAGGCTTAACGGTCAAAATATCAATTGAAACACTATTCCTGT 179		
DB	718 AGATGGCTTATCACTTACTCGGGAGTACGGGATATTTCAAATTGAAAGTTTTTTAAGTAA 777		
QY	180 AAGTAATGAAATCGTGGTTTGCTTAATTAATCTCTCCAGGAGCTATTTTATTAGAGCTAG 239		
DB	778 AGATGATGAAAAAGCTCAAAATGACCATTTGAACGCGCAGGTAGCATCGTTTTTACAATCTCG 837		
QY	240 TTTTATTTATTAAATTATTATTTTCAAGTTTGGCAGATATTAGTATAAATCTTAAAGAAATTTGA 299		
DB	838 TTTCTTTGGCGAAATTATTCGTAATTAACAGAGATATGTTCAATGCGAAGTTCTAGA 897		
QY	300 ACAACACCAAGTTGTTTTTAACAGTGTGTAATACAGAGATTACCTTAAAAAGGAAAAGATGT 359		
DB	898 TAACAATCAAGTTGCAATTACTTCAGGAAAAGCTGATTTTACGGTTAATGGAATTAGATGC 957		
QY	360 TGACCAAGTATCCCTGTCTACAGAGATATCAACAGAAATTCCTTTGATTTTAAAAACAAA 419		
DB	958 TGAATAATTATCCGCACCTTACCGTAAATTGATCTCAAAACCAATGAAATTAATCTGTTC 1017		
QY	420 ATTATTGAAGTCTATTATTGCTGAAACAGCTTTTTCAGCCAGCTTTTACAGAAAGTTCGTCC 479		
DB	1018 TTTATTGACAAAATTAATTAGTGAACAGGTTTTCGTGATTCGATGACGAAAGTTCGTCC 1077		
QY	480 TATTTTAAACAGGAGTTCATATTGTTATTAGTAATCATATAAGATTTTAAAGCAGTAGGCAC 539		

XX New isolated nucleic acid derived from *Enterococcus faecium* encoding an
PT *Enterococcus faecium* polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
XX Example 1; SEQ ID NO 902; 243pp; English.
XX

CC The invention relates to an isolated nucleic acid derived from
CC *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids are useful for diagnosing pathological conditions
CC resulting from *E. faecium* bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of *Candida albicans*.
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating *Enterococcus faecium* infections. The present sequence represents
CC one of the disclosed *E. faecium* nucleic acids.
XX

XX Sequence 1155 BP; 369 A; 215 C; 215 G; 356 T; 0 U; 0 Other;

Query Match 28.0%; Score 317; DB 10; Length 1155;

Best Local Similarity 55.6%; Pred. No. 4.7e-58;

Matches 630; Conservative 0; Mismatches 500; Indels 3; Gaps 1;

QY 1 ATGATTCAATTTTCAATTAATCGACATTAATTTTATTTATTCATGCTTTTAAATACAACTAAAGCT 60
DB 22 ATCATGAAGTTTACITTAACCGAGCTAGCTTTATGCGAGGATTCGAACCTGTTCAACGA 81
QY 61 GCTATTAGCACTAAATGCGATTCCTATTTCTTTCATCAATAAATGAAAGTCACTTCT 120
DB 82 GCTATTCAAGCAAAACCAACGATCCCTATTTTGACAGGTGTAATAATCACACTGACACAA 141
QY 121 ACAGGAGTAACCTTTTAAACAGGCTTAACGCTCAATATCAATCAATGAAACACTATTCCTGTA 180
DB 142 GAGGTTTGACTTTGACGGGGGAGCAACGCTGATATATCAATTTGAACTTTTGTGCTGT 201
QY 181 AGTAAGAAATGCTGGTTGCTAATATCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
DB 202 GAAACGAAAGCAAAATATGCAATCGAATCTACTGTTCCATTTGTTTACAGCAGCT 261
QY 241 TTTTATTATTAATTTTCAAGTTTGGCAGATATAGTATAAATGTTTAAAGAAATGAA 300
DB 262 TCTTTAGCGAAATCAATTCGGAGACTTCTCGAAGAACATTTACTTTAGAAGTTTAA 321
QY 301 CAACACCAAGTTGTTTAAACAGGCTGTAATCAAGATTTACTTTAAAGGAAAGATGTT 360
DB 322 AATAACCAAGTAGCGATCACTTCGAAAGAGGATTTTATCGTAATGATAGTAGCA 381
QY 361 GACGAGTATCTCGTCTCAAGAGATATCAACAGAAATCTTTGATTTTAAACAAAA 420
DB 382 GATAACTATCTCTATCTTCTGTTGCGAAAGCCATAACAGATCAAAATACCTGTACAC 441
QY 421 TTATTGAAGTCTATTATGCTGAAACAGCTTTTCGACCGAGTTTACAGAAAGTCGCTCT 480
DB 442 GTATTGACTAACTAATCAAGAAACAGTTTTCGTTCTTCAACATGAGAGTCGTTCCA 501
QY 481 ATTTTACAGAGTTTCAATTTGATTAAGTAAATCAATAAGATTTTAAAGCAGTAGCGACT 540
DB 502 ATCTTGACAGGTGTCATTTTCAATTTA--TCTGATAATCTTTTATAGTGTAGTACT 558
QY 541 GACTCTCATCTATGACGACAGCTTTTATCACTTTGGCAATATCTCAGCAGATTTGATG 600
DB 559 GATTCTCACCCTTAAGTCAACGCGTGAATTCAGTAGAACAAGCGGCTGATCATTTTGTAT 618

QY 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATATTGAG 660
DB 619 ATTGTTATTCCTGGAAAAAGTTTGATCGAATTTATCTCGCTCATTAACAAATGAAGAGAA 678
QY 661 ACCGTTGAGGTATTTTCTCACCAAGGCCAAATCTTTGTTTCAAGAGTGAACACATTTCTTTT 720
DB 679 ATCGTCGAAATCAGCATTTATGGAACACCAAGTGTGTTCAAAACAGAACGATGATTTTC 738
QY 721 TATACACCCCTCTTAGAAGAGAAATTTATCCCATACAGACCGTTTATTAATGACAGAAATTT 780
DB 739 TATTCCCGCTTTGCTAGAAGAGAACTATCCAGATACCAATCGTTTGTGATTTCTTCAAGTTT 798
QY 781 GAGACGAGGTGTTTTCATACCAATCCCTTCGACAGCTATGGAACGTCCTTCTTTG 840
DB 799 AATACAGAGGTGTAATTTTCTGTTCTTCTAGCTTTTTCGCGGATCGAACCGGCTTCTTTA 858
QY 841 ATTTCTAATGCTACTCAAAATGTTTAAAGCTTCAGATTTACTCAAAATCATATTTCA 900
DB 859 CTTTCTCATGAAGCGCGTAACACATCGTTTCTGTTTATCTTATTCGTCAGATCGGTTGTT 918
QY 901 GCTCATGTTAACTCACTGAGGTGTTAGTAACAGAGATTTAGATATTGTTAGTCAG 960
DB 919 TTATACGAAATTCGCGAGAAATCGGAAAGTAGAGAAAGCTTGAGTTATACAGCAAGC 978
QY 961 TCTGTTAGTATTTAACTATTCAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 1020
DB 979 AGTGGCGATCCATTAATGATTTTCTTCAATCCGATTTATGAAGCAGCGTTCCGTGCA 1038
QY 1021 ATTTAAAGTGAACAGATTAATAATTTCTTATCACCAGTTTCGACCAATTCACCTAACA 1080
DB 1039 TTTGAGATATGAGATCAAAAGTGAATTTATCTCTGCTATTCGTCATTTACATTGAA 1098
QY 1081 CCAGCGCATGAGGAAGAAAGTTTATCCAAATTAATTAACAGATACGAAACAA 1133
DB 1099 CCAACGGAAGATCGGTTCAATTTTATTTAGTCTTCTGATACGATACAA 1151

RESULT 21

ACA33953

ID ACA33953 standard; DNA; 1128 BP.

XX ACA33953;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #15610.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX *Enterococcus faecium*.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342933P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU30083.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
PS Claim 14; SEQ ID NO 21823; 1766pp; English.

XXXX

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences

Sequence 1128 BP: 355 A: 210 C: 212 G: 351 T: 0 U: 0 Other:

Query Match	27.8%	Score 315.6;	DB 8;	Length 1128;
Best Local Similarity	55.6%	Pred. No. 9.2e-58;		
Matches 628;	Conservative	0;	Mismatches 499;	Indels 1;
				Gaps 1;

Qy	4	ATTCAAATTTTCAATTAATCGCACATATTATTATTCATGCTTTAAATACAACTAAACGTC	63
Db	1	ATGAAGTTACTTTAAACCGAGCTAGCTTTATGCAGGAATTCGAAACTGTTTCAACGAGCT	60
Qy	64	ATTAGCACTAAAAATGCAATTCCTATTCTTTTCATCAATAAAATGGAAGTCACITCTAC	123
Db	61	ATTTCAAGCAAAACACGATCCCTATTATTTGACGGTGTAAAAATCACACTGACACAAGAA	120
Qy	124	GGAGTAACTTTAACAGGGTCTAAACGGTCAAATATCAATTTGAAACACATATTTCCTGTAAGT	183
Db	121	GGTTTGACTTTGACGGGAGCAACGCTGATATCAATTGAACCTTTTGTCTGTGTGAA	180
Qy	184	AATGAAATATGCTGTTTGTCTAAATTAACCTCCAGGAGCTATTTATTAGAAGCTAGTTTT	243
Db	181	AACGAAAAAGCAATATGCAAAATCGAATCTACTGTTCCATTGTTTACAAGCAGCTTTC	240
Qy	244	TTTATTAAATATTATTTCAAGTTTTCGACATATTAGTATAAATGTTAAAGAAATTTGAACAA	303
Db	241	TTTAGCGAAATCATTTGCGAGACTTCCCTGAAGAAACATTACTTTAGAGTTTGTAGAAAT	300
Qy	304	CACCAAGTTGTTTTAAACAGTGGTAAATACAGATTAACCTTAAAGGAAAGAGTTGAC	363
Db	301	AAAACAAGTAGGCATCACCTTCGAAAGCGAAATTTTATCGTAAATGGATTAGATGCAGAT	360
Qy	364	CAGTATCCTCGTCTACAGAAGATATCAACAGAAATCCTTTGATTTTAAAAACAAATTA	423
Db	361	AACATATCTCATCTCTGTTGTGCGAAGCCATACACAGATGAATTTACCTGTACACGTA	420
Qy	424	TTGAAGTCTATTATTGCTTGAAACAGCTTTTTCAGCCAGTTTTCAGGAAGTCGTCCTATT	483

XX 30-APR-2002.
PD XX
PF 13-AUG-1998; 98US-00134001.
XX
PR 14-AUG-1997; 97US-0055779P.
PR 08-NOV-1997; 97US-0064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Doucette-Stamm LA, Bush D;
PI
XX
XX WPI; 2002-381255/41.
DR P-PSDB; ABP40321.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX Disclosure; SEQ ID NO 2329; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
XX Sequence 1158 BP; 432 A; 177 C; 192 G; 357 T; 0 U; 0 Other;
SQ
Query Match 21.8%; Score 247.4; DB 6; Length 1158;
Best Local Similarity 52.9%; Pred. No. 3 3e-43;
Matches 580; Conservative 0; Mismatches 511; Indels 6; Gaps 2;
QY 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 25 ATGATGGAATTCACAAATTAAGAGATTTATTTTATTAATCAACTTTAAACGACATTTAAAG 84
QY 61 GCTATTAGCATAAATGCGCATTTCTTCTTTCATCAATTAATTAATGAGTCACTTCT 120
Db 85 GCTATCTCACCTAGAACAACTTTACCGATTTTAAACAGGTATCAAAATTTGATGCTAAAGAA 144
QY 121 ACAGGAGTAACTTTAAACAGGGTCTAACCGTCAATATCAATTTGAAACACATTTCC--T 177
Db 145 AATGAAGTCATTTTAACTGGGTGAGTTTCAAGAAATATCAATAGAAATCACTATTCAAA 204
QY 178 GTAAGTAATGAATGCTGGTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 237
Db 205 CAAGTTGATGGTGAAGAAATTTTGAATTTACAGAAACAGGATCAGTTGTTCTCTGCT 264
QY 238 AGTTTATTTATTAATTAATTTTCAAGTTTGGCAGATATAGTATTAATTAATTAAGAAAT 297
Db 265 CGTTCTTCTGTTGATTTATTAATTAATTAATTTCTCGAAAGAAAGTTTAAATTTATCACT 324
QY 298 GAACAAACACCAAGTTGTTTAAACAGGTGTTAAATCAGAGATTACCTTAAAGGAAAGAT 357
Db 325 GAACAAATTCACAACTTAATCAATCAGGTGATTTTAACTTAAGTGGCTTAGAT 384
QY 358 GTTGACCAAGTATCTCGTCTACAAGAGATATCAACAGAAATTCCTTTGATTTTAAAGAA 417
Db 385 CCTGATCAATACCCATTTATTAATGAGTATCAGGATGAGCGTATTCAATTTGTCAGTT 444
QY 418 AATATTATTAAGTCTATTTATTTGCTGAAACAGCTTTTGGACGAGTTTACAAGAAAGTCT 477
Db 445 AAGGTACTAAAAAATATATTCAGCAAACTAAATTTTTCAGTGTCCACCTCAGAAACACGA 504
QY 478 CCTATTTTAAACAGGAGTTCATATTTGATTAATTAATTAATTAATTAATTAATTAATTA 537
Db 505 CCAGTACTTACTTGGTGAACCTGGCTTATA--CAAGATTAATGAATTAATTAATTAATTA 561

QY 538 ACTGACTCTCATCGTATGAGCCAAAGTTTAAATCACTTTTGGACAATACTTTAGCAGATTTG 597
Db 562 ACAGATTTCACACCCGCTTAGCTGTAAAGAAAGTTACAGTTTAGAAGATGAATCAGAAATAA 621
QY 598 ATGGTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGAT 657
Db 622 AATGTCATCTCTCTGTTAAAGCTTTTATCTGAAATTAACAAATTTATGAGTGACAGCGAC 681
QY 658 GAGACCGTTGAGGTATTTTCTCAACCAAGCCAAATCTTTTTCAGAGTGAACACATTTCT 717
Db 682 GAAGATATTGATATTTTCTTCTTCTTAAACCAAGTGTATTTCAGAGTGGGAAATATTAA 741
QY 718 TTTTATACACCCCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAA 777
Db 742 TTTATCTCACGTTTACTTTGAAGGTCATTTATCCAGATACGACACGTTTATTTCCAGANA 801
QY 778 TTTGAGACGAGGTGTTTCTCAATACCAATCCCTTCCGACGCTTATGGAAGTGCCTTTC 837
Db 802 TATGAGATTAATTAGGAAATTAACATGGAGACTTCTATCATGCAATTTGATCGTCATCT 861
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAAGCTTTGAGATTTACTCAAAATCATAT 897
Db 862 TTATTAGCACGTTGAAGGTGAAATAAATGTTTAAATTAAGTACAGGTAATGAATTAGTT 921
QY 898 TCAGTCTATGTTAACTCACCTGAGGTGGTAAAGTAAACGAGGATTTAGATATTTGTTAGT 957
Db 922 GAACCTTTCATCTACTTCTCTGAAATTTGGTACTGTTTAAAGAAAGATTTAACGCTAATGAT 981
QY 958 CAGTCTGTAGTGTATTTAACTATCATGCTTTCAATCAACTTACCTTATTTGAGTCTTTAAAA 1017
Db 982 GTAGAAGCGGAAACTTGAATAATTTCTTTCACTCAAAATACATGATGATGCTTTAAAA 1041
QY 1018 GCTATTTAAAGTGAAACAGTAAATAATTTCTTTTATCACCAGTTTCGACCATTTACCCCTA 1077
Db 1042 GCCATTGATATGATGAAGTAGAAGTAGAATTTCTTTGGTACAATGAAACCATTTATCTTA 1101
QY 1078 ACACCAGCGATGAGGA 1094
Db 1102 AACCCAAAGATGATGA 1118
RESULT 24
ACAT47854
ID ACAT47854 standard; DNA; 1131 BP.
XX ACAT47854;
XX AC
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #29511.
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Staphylococcus haemolyticus.
OS
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX

WP ADF77343_18 1800001 1910000
 WP ADF77343_19 1900001 1983043
 ID ADF77343 standard; DNA; 1983043 BP.
 XX
 AC ADF77343;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Lactic acid bacteria *Lactobacillus johnsonii* Lal genomic DNA SEQ ID NO:1.
 XX
 DE Lactic acid bacterium; *Lactobacillus johnsonii* strain Lal; detection;
 KW identification; DNA array; DNA chip; gene; ds.
 XX
 OS *Lactobacillus johnsonii*.
 OS
 PN WO2003084989-A2.
 XX
 PD 16-OCT-2003.
 XX
 PF 19-MAR-2003; 2003WO-EP002882.
 XX
 PR 09-APR-2002; 2002EP-00007932.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 XX Pridmore RD, Mollet B, Arigoni F, Hermanns J;
 PI
 XX WPI; 2003-804292/75.
 DR
 XX
 XX Use of a DNA sequence of a *Lactobacillus johnsonii* strain, Lal, or its
 PT part or sequences homologous to it for elucidating interactions between a
 PT bacterium and host that it colonizes.
 PS
 XX Claim 1; SEQ ID NO 1; 24pp; English.
 CC
 CC The present sequence represents a DNA sequence of the lactic acid
 CC bacteria *Lactobacillus johnsonii* strain Lal. The present invention also
 CC described the DNA sequence, or its part or sequences homologous to it
 CC which are useful for elucidating interactions between a bacterium and
 CC host that it colonizes. Also described: (1) a method for detecting,
 CC identifying or selecting a *Lactobacillus* strain, preferably *Lactobacillus*
 CC *johnsonii* in a biological sample; (2) a method for detecting, identifying
 CC or selecting antibodies directed to *Lactobacillus* polypeptides in a
 CC biological sample; (3) a DNA array or chip containing an array of
 CC polynucleotides comprising at least a polynucleotide derived from Lal;
 CC (4) a protein array or chip containing an array of polypeptides
 CC comprising at least one of the polypeptides obtainable by expressing a
 CC polypeptide as identified by an open reading frame derived from Lal; (5)
 CC an antibody chip containing an array of antibodies comprising at least
 CC one antibody directed to a polypeptide obtainable by expressing an open
 CC reading frame in Lal; (6) a screening assay; (7) a kit comprising the
 CC polynucleotide or antibody raised against the polypeptide; (8) a computer
 CC readable medium having recorded the nucleic acid sequence of Lal or the
 CC polypeptide derived from the nucleotide sequence; and (9) a computer-
 CC based system for identifying fragments of the *Lactobacillus johnsonii*
 CC genome. The DNA sequence of a *Lactobacillus johnsonii* strain Lal, or its
 CC part or sequences homologous to it are useful for elucidating
 CC interactions between a bacterium and host that it colonizes.
 XX
 SQ Sequence 1983043 BP; 655813A; 330432C; 354497G; 642272T; 0U; 290Other;
 Query Match 21.3%; Score 242; DB 10; Length 110000;
 Best Local Similarity 53.4%; Pred. No. 1e-41;
 Matches 609; Conservative 0; Mismatches 510; Indels 21; Gaps 4;
 QY 1 ATGATTCAATTTTCAATTAATCGCATATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
 DB 1543 ACGATGCGATTTACAAATTAATCGTAAATTTTCTCGAAACCTAAATTAATGCTATGCGT 1602
 QY 61 GCTATTAGCAGTAAATTAATCGCATATTTATTTCTTCATCAATAAATAATGAAGTCACTTCT 120
 DB 1603 GCAATTTCTTACGCTGCTACTATTTCCATATTAAGTGGTATTAATAACTTACCCCTTACAGT 1662

QY 121 ACAGAGTAACCTTTAAACAGGGTCTTAACGGTCAATATATCAATTAATGAACACATATTCCTGTA 180
 DB 1663 GAAATGCTTACTTTAAACCGGTAGTACTGACATTTCTTATTGAATTTCAATCCCACT- 1721
 QY 181 AGTAATGAAATGCTGGTTTCTAATACCTCTCCAGGAGCTATTTTATTATAGAGCTAGT 240
 DB 1722 -----AAACGATGATTTAATTTGTTCAATCTACGGATCGATCGTTTTTACCACAGA 1773
 QY 241 TTTTATTATTAATTTTCAAGTTTGCAGATATTTAGTATATAAATGTTTAAAGAAATTGAA 300
 DB 1774 TTTTTCAGTGAATTTGTAATAAATAATACCAGTAAAGATTTTTCATTTGAAAGTAAAGAA 1833
 QY 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAGATGTT 360
 DB 1834 AGTTTTCAAACGAAAAATTTGTTTCTGAAATACCGAATTCATGATTAACGGTTTAGATGA 1893
 QY 361 GACCAAGTATCTCGTCTACAAAGAGATATCAACAGAAAAATCTTTGATTTTAAACAAAA 420
 DB 1894 AACAAATTAATCCACTTACAGAAATTTCTACTGATGATCATTTCAAAATTTCTGGTAAA 1953
 QY 421 TTATTGAAGTCTATTTTGTGAAACAGCTTTTGCAGCCAGTTTTCAGCAGTATTTCAAGAAAGTCTCT 480
 DB 1954 ACATTTAGAGAAATTTATTAACGAAATCTTTTGCAGTCTGCTACTCAAGAAAGTCTGCCA 2013
 QY 481 ATTTTACAGAGATTCATATTTGATTAAGTAAATCATAAAGATTTTAAAGCAGTAGCGACT 540
 DB 2014 ACTTTAACTGGTAACTTTATCTT---CAACAATTCATCAATTAAGCAGTTGCTTACC 2070
 QY 541 GACTCTCATCTGATGAGCCAACTTTAATCACTTTGGACAATA---CTTCAGCAGATTTG 597
 DB 2071 GACAGTACCAGTATCTCAACGTCAAATTTCTCTAGAAATGGTCTTCAACAAGTACT 2130
 QY 598 ATGGTAGTTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATAT 657
 DB 2131 GACTTAATTTATCTCGGAAAGAGCTTAGTAGAATTTATCCCGAATTTATCGGAGAAAGTAT 2190
 QY 658 GAGACGGTGGAGTATTTTCTCCACAGCCAAATCTTTGTCAGNAGTGAACACATTTCT 717
 DB 2191 CCTGAAATCACAGTAAATCTGGTGGAAACCAAGTTTATTTGAAAGTTGGAACATGCA 2250
 QY 718 TTTTATACACCCCTCTTTAGAGAGAAATTTATCCCATACAGACCGTTTATTAATGACAGAA 777
 DB 2251 TTCTATTACGTTTACTTGTAGGACATATCCAGACACTGATCGTTTAATTTCCAATGAA 2310
 QY 778 TTGAGACGGAGTGTGTTTCAATACCAATCCCTTCGCCACGCTATCGAACGTCGCTTC 837
 DB 2311 TCTACAACATCAGTTGAATTTGAAATTTACAGTCTTTTAGCTCTCTCTCGAACGTCGCACT 2370
 QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAG-----CTTGAGATTTACTCAAAAT 891
 DB 2371 CTTCTTACTACGAAAGCCGTAATTAATGTTGTTAAGATGACTCTTTGATTTTCAAAATCAA 2430
 QY 892 CATATTTTCAGCTCATGTTTAACTCACTCAGGTGGTAAAGTAAACGAGGATTTTAGATAT 951
 DB 2431 TTAGTTAACTCCAGGTGATTTACCGGAATAGGAATAGGAATGTTGGAAGAGAAATTTGGCTTT 2490
 QY 952 GTTAGTCTAGTCTGGTAGTATTTAACTATCAGCTTCAATCCAACCTTACTTATTGAGTCT 1011
 DB 2491 AAGAATCTTGAAGGAGATGGCTTAATTTCTTCAATCTCTGATTTATTAAGAGAGAGCT 2550
 QY 1012 TTAAGAGTATTAAGTGAACAGTAAATTTCTTTTATCACCAGTTTCGACCATTC 1071
 DB 2551 TTACGTGCTCTATTTACTGATTTCCATTTATTAAGATTTTACCAGCCGCTAAGACATTT 2610
 QY 1072 ACCCTAACACCCAGCGCATGAGAGAAAGTATTTTATCCAAATTAATTAACACCAAGTACGAACA 1131
 DB 2611 ACAGTGATACCGCAACAAAGATGTTTAATTTCTCAATTAATCAACCCAGTTAGAACACA 2670
 RESULT 26
 ACF74365
 ID ACF74365 standard; DNA; 1131 BP.
 XX

AC	ACF74365;	QY	358	GTGACGATATCCCTCGTCTACAGAGATATCAACAGAAATCCCTTTGATTTTAAAAACA	417
XX		Db	361	CCAGATCAATATCCCTTTATTTACCTCAAGTTTCTAGAGATGACGCAATTCATTTGTCGTA	420
XX	20-NOV-2003 (first entry)	QY	418	AAATATTGAAGTCTATTTATTTGCTGAAACAGCTTTTGACGACGTTTCAACAGAAAGTCGT	477
XX	Staphylococcus aureus DNA #2045.	Db	421	AAAGTGTGTTAAACAGCTGATTTGCACAAACAAATTTGCGAGTGTCCACCTCAGAAACACGC	480
XX	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;	QY	478	CCTATTTTAAACAGAGATTCATATTTGTTAAGTAAATCAATAAGATTTTAAAGCAGTAGCG	537
KW	enzymatic assay; antibiotic target; gene; ds.	Db	481	CCAGTACTAACTGGTGTGAACCTGCTTATA---CAAGAAATGAATTAATATGCACACGC	537
XX	Staphylococcus aureus.	QY	538	ACTGACTCTCATCGTATGAGCCCAAGTTTAATCACTTTTGGACATATCTTCAGCAGATTGG	597
XX	WO200294868-A2.	Db	538	ACTGACTCACACCGCTTGGCTGTGAAGATTTGACAGTTAGAAAGATGTTCTTGAACACAA	597
XX	28-NOV-2002.	QY	598	ATGGTAGTTCCTCAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATATT	657
XX	27-MAR-2002; 2002WO-IB002637.	Db	598	AATGTCATCATTTCCAGGTAAAGCTTTAGCTGAATTAATAAATTAATGCTCTGACATGAA	657
XX	27-MAR-2001; 2001GB-00007661.	QY	658	GAGACCGTTGAGGTATTTTCTCACCAGCCCAATCTTTGTCAGAAAGTGAACACATTTCT	717
XX	(CHIR-) CHIRON SPA.	Db	658	GAAGACATTTGATATCTTTCTTTGCTTCAACCCAAAGTTTATTTAAAGTTGGAATGTGAAC	717
XX	Masignani V, Mora M, Scarselli M;	QY	718	TTTTATACACGCTCTTTAGAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAA	777
XX	WPI: 2003-120786/11.	Db	718	TTTTATTTCTGATTTATAGAGGACATTTCTGATACACACGTTTATTTCCCTGAAAC	777
XX	P-PSDB; ABW72805.	QY	778	TTTGAGCGGAGGTGTTTTCATATACCAATCCCTTCGCCACGCTATGGAAACGTCCTTC	837
XX	New Staphylococcus aureus protein, useful as a vaccine for treating or	Db	778	TATGAAATTAATTAAGTATAGACATGGGAGTTTATCATGCGATTGATCGTCTCT	837
XX	preventing Staphylococcal infection, specifically an infection caused by	QY	838	TTGATTTTCTAATGCTACTCAAAATGGTACTTGAAGCTTGAGATTACTCAAAATCATATT	897
XX	S. aureus, e.g. sepsis.	Db	838	TTATAGCGGTGAAGGTGATACGTTATTAATTAAGTACAGGTGATGACGTTGTT	897
XX	Claim 6; SEQ ID NO 4089; 49pp; English.	QY	898	TCAGCTCATGTTTAACTCACCTGAGGTGGTAAAGTAAACGAGGATTTAGATATTGTTAGT	957
XX	The invention relates to novel genes and encoded proteins from	Db	898	GAATTTGCTTCTACATCACACAGAAATTTGGTACTGTAAGAGAGAGATTTGATGCAACGAT	957
XX	Staphylococcus aureus. A composition comprising the S. aureus protein, a	QY	958	CAGTCTGGTAGTATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTAAAA	1017
XX	nucleic acid encoding the protein, or an antibody to the protein, is	Db	958	GTTGAAGGTGGTAGCCTGAAATTTTCATCAACTCTAAATATATATGATGATGATGCTTTAAA	1017
XX	useful as a pharmaceutical, particularly as a vaccine for treating or	QY	1018	GCTATTTAAAGTGAACAGATAAATTCATTTCTTATCACAGTTCGACCATTCACCCCTA	1077
XX	preventing infection due to Staphylococcus bacteria, specifically an	Db	1018	GCAATCGATATGATGAGTGTGAAGTTGAATTTCTTCGTAACAATGAACCATTTATTCTA	1077
XX	infection caused by S. aureus. The composition is particularly useful for	QY	1078	ACACCGGCGATGAGGA 1094	
XX	treating or preventing sepsis in a patient. The composition can also be	Db	1078	AAACCAAAGGTGACGA 1094	
XX	used for diagnostics. The protein is also used in an assay for enzymatic				
XX	studies and as a target for antibiotics. This sequence represents one of				
XX	the novel S. aureus genes of the invention				
XX	Sequence 1131 BP; 408 A; 177 C; 195 G; 351 T; 0 U; 0 Other;				
XX	Query Match 20.8%; Score 236.2; DB 8; Length 1131;				
XX	Best Local Similarity 52.2%; Pred. No. 8.1e-41;				
XX	Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;				
QY	1 ATGATTCAATTTCAATTAATCGCACATTATTTATTTATTCATGCTTTAAATACAACTAAACGT 60				
Db	1 ATGATGGAAATTCACATTAAAGAGATTTATTTATACCAATTAATGACACATTAAAA 60				
QY	61 GCTATTAGCATTAAATGCCATTCCTATTTCTTTCATCAATAAAAAATGAAGTCACTTCT 120				
Db	61 GCTATTTCCACCAAGAACCAATACCTATTAATTAATGATGATCAAAATCGATCGAAGAA 120				
QY	121 ACAGAGATTAATTTAAAGGTCTTAACCGTCAATATCAATATGAAGAAACATTTTCTGTA 180				
Db	121 CATGAAGTATTAATTAATGCTGATTTTCAAGTCTGAAATTTCAATAGAAATCACTATTCTTAA 180				
QY	181 AGT---AATGAAATGCTGGTTGCTAATTAATACCTCTCCAGGAGCTATTTATTTAGAGCT 237				
Db	181 ACTGTAGATGCGGAGATATTGTCATATTTTTCAGAAACAGGCTCAGTAGTACTTCTTGGGA 240				
QY	238 AGTTTTTTTATATATTTTCAAGTTTGGCAGATATTTAGTATAAATGTTTAAAGAAAT 297				
Db	241 CGATTTCTTTGATATTAATAAAAAATTTACCTGCTAAGATGTTTAAATTTATCTACAAAT 300				
QY	298 GAACAAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGAT 357				
Db	301 GAACAAATCCAGACATTAATTAATACATCAGGTCAATCTGAAATTTAATTTAAGTGGCTTAGAT 360				
XX	RESULT 27				
XX	AAS4991				
XX	ID AAS54991 standard; DNA; 1134 BP.				
XX	AC AC				
XX	AAS54991;				
XX	13-FEB-2002 (first entry)				
XX	Staphylococcus aureus DNA for cellular proliferation protein #1303.				
XX	Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;				
XX	antibacterial; drug design.				
XX	Staphylococcus aureus.				
XX	WO200170955-A2.				
XX	27-SEP-2001.				
XX	21-MAR-2001; 2001WO-US009180.				

30-NOV-2001; 2001WO-CA001754.	XX
30-NOV-2001; 2000US-00727892.	XX
20-JUN-2001; 2001US-00885561.	XX
(PHAG-) PHAGETECH INC.	XX
Pelletier J, Gros P, Dubow M;	XX
WI; 2002-566592/60.	XX
P-PSDB; ABB81847.	XX
Use of a <i>Staphylococcus aureus</i> (STAAUR2) polypeptide or composition comprising STAAUR2 and polypeptide encoded by bacteriophage open reading frame that specifically interacts with STAAUR2, for identifying a compound active on STAAUR2.	XX
Disclosure; Fig 1; 131pp; English.	XX
The invention relates to a novel <i>Staphylococcus aureus</i> polypeptide (STAAUR2), and its use in drug screening assays. The polypeptide of the invention has antibacterial activity. The compound active on STAAUR2 is useful for the manufacture of an antibacterial agent or for the manufacture for medication for treating or preventing a bacterial infection. The sequence encodes the <i>S. aureus</i> STAAUR2 of the invention	XX
Sequence 1134 BP; 410 A; 177 C; 195 G; 352 T; 0 U; 0 Other;	XX
Query Match	20.8%; Score 236.2; DB 6; Length 1134;
Best Local Similarity	52.2%; Pred. No. 8.1e-41;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;	
1 ATGATTCATTTTCAATTAATCGCACATATTTATTTATTCATGCTTTAAATACAACTAAAGCT 60	QY
1 ATGATGGAATTCATTTAAAGAGATTTATTTTATACAAATTAATGACACATTAATAA 60	Db
61 GCTATTAGCACTAAAAATGCCATTCCTATCTTTCATCAATAAAAAATGAAAGTCACITCT 120	QY
61 GCTATTTTCCACGAACAACATTCCTATATTAACTGGTATCAAAATCGATCGGAAGAA 120	Db
121 ACAGGAGTAACTTTACAGGCTCAACGGTCAAAATCAATGAAACACATTTCTCTGTA 180	QY
121 CATGAAGTATATTAATCTGGTTCAGACTCGAAATTTCAATAGAAATCACTATTCTTAAA 180	Db
181 AGT---AATGAAATGCTGGTTGCTAATTTACTCTCCAGGAGCTATTTTATTAGAAGCT 237	QY
181 ACTGTAGATGGCGAAGATATTGCTCAATATTTTCAGAAACAGGCTCAGTAGTACTTCTGGA 240	Db
238 AGTTTTTTTATTAATTAATTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAATT 297	QY
241 CGAATCTTTGTTGATTTATAAAAAATTTACTCGTAAAGATGTTAAATTTATCTACAAAT 300	Db
298 GAACACACAAAGTTGTTTAAACCAAGTGGTAAATCAGAGATTACTCTTAAAGGAAAGAT 357	QY
301 GAACAAATCCAGACATTAATATACATCAGGTCATCTCGAATTTAATTAAGTGGCTTAGAT 360	Db
358 GTTGACCAATCTCTGCTACAGAAGTATCAACAGAAATTCCTTTGATTTTAAACA 417	QY
361 CCAGATCAATATCTTTTATCTCAAGTTTCTAGAGATGACGCAATTTCAATTTGTCGTA 420	Db
418 AAATATTGAAGTCTATATTGCTCAACACAGCTTTTTCAGGACAGTTTACAGAAAGTCGT 477	QY
421 AAAGTCTTAAACCGTGAATGCACAAACAAATTTTCAGTGTCCACCTCAGAAACACGC 480	Db
478 CCTATTTTAAAGAGTTCATATGTTATAGTAAATCAATAAGATTTTAAAGCAGTAGCG 537	QY
481 CCAGTACTAACTGGTGTGAATGCGCTTATA---CAAGAAATGAATTAATATGACACGC 537	Db
538 ACTGACTCTCATCGTATGAGGCAAGTTTAACTACTTTGGAACATATCTCAGCAGATTG 597	QY
538 ACTGACTCACCGCTTGGCTGTAGAAGTTGAGTTAGAAATGTTTCTTCTGAAACAAA 597	Db


```

PR 04-OCT-2001; 2001US-0327193P.
PR 30-OCT-2001; 2001US-0340922P.
PR 05-NOV-2001; 2001US-0338709P.
PR 06-NOV-2001; 2001US-0333269P.
PR 18-DEC-2001; 2001US-0341679P.
XX
PA (APFI-) AFFINIUM PHARM INC.
XX
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowsmith C, Awrey D;
PI Beattie B, Canadien V, Cox B, Domagala M, Houston S, Li Q;
PI Nethery K, Ng I, Ouyang H, Pinder B, Sheldrick B, Viola C;
PI Wrezel O;
XX
XX WPI; 2003-371793/35.
DR P-PSDB; ADD26272.
XX
XX New crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae, Helicobacter pylori, Pseudomonas aeruginosa
PT involved in nucleic acid processing, useful as targets for pathogenic
PT bacteria.
XX
XX Claim 218; SEQ ID NO 76; 298pp; English.
XX
XX The present invention describes a crystallised recombinant polypeptide
CC (I) comprising the amino acid sequence of polypeptides from
CC Staphylococcus aureus, Streptococcus pneumoniae, Helicobacter pylori and
CC Pseudomonas aeruginosa which are involved in nucleic acid processing, or
CC amino acid sequences having at least 90 % identity with the polypeptide
CC sequence, where the polypeptide is in crystal form. (I) comprises the
CC amino acid sequence (S) of polypeptides involved in nucleic acid
CC processing, which includes ribonuclease diphosphate reductase, major
CC subunit (nrDE), polyribonucleotide phosphorylase (pnppA), transcription
CC termination factor rho (Rho), putative polynucleotide
CC nucleotidyltransferase (pnp), DNA ligase (lig), DNA polymerase III, beta-
CC subunit (dnan) or ribonucleoside diphosphate reductase minor subunit (R2)
CC (nrdF) from S. aureus; uridylylate kinase (pyrH) or orotate
CC phosphoribosyltransferase (pyrE) from H. pylori; uracil-DNA-glycosylase
CC (lung) from P. aeruginosa; and nrdE from S. pneumoniae. (I) is in a
CC crystal form. (I) has antibacterial activity, and can be used in
CC vaccines. (I) is useful for designing a modulator for the prevention or
CC treatment of S. aureus, S. pneumoniae, H. pylori and P. aeruginosa
CC related diseases or disorders. The method comprises providing a three-
CC dimensional (3D) structure for (I), identifying a potential modulator by
CC reference to the 3D structure, contacting the potential modulator with
CC the recombinant polypeptide and assaying the activity of the polypeptide
CC or determining the viability of S. aureus, S. pneumoniae, H. pylori or P.
CC aeruginosa after contact with the modulator, where a change in the
CC activity of the polypeptide or the viability of the bacteria indicates
CC that the modulator may be useful for preventing or treating the disease
CC or disorder. The structural and functional information of (I) aid in the
CC discovery and design of therapeutic and diagnostic molecules. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 1134 BP; 410 A; 177 C; 195 G; 352 T; 0 U; 0 Other;

Query Match 20.8%; Score 236.2; DB 10; Length 1134;
Best Local Similarity 52.2%; Pred. No. 8.1e-41;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTCAATTTTCAATTAATCGCATTATTTATTTATTCATCAATAAATAATGAAAGTCACTTCT 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 ATGATGGAAATTCACATTAATAAGAGATTATTTATTTATACAAATTAATGACATTAATAA 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCTATTAGCAGCTAAATAATGCCATTCTCTATTTCTTCATCAATAAATAATGAAAGTCACTTCT 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCTATTTCCCAAGAACCAACATTACCTATATTAATCTGTTATCAAAATCGATCGGAAGAA 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ACAGAGTAACCTTTAAACGGGCTAAACCGGTCAAAATATCAATTTGAAAAACATTTCTCTGTA 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CATGAAGTTATATTAACCTGGTTTCAGACTCTGAAATTTCAATAGAAATCACTATTCTCTAAA 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 AGT---AATGAAATGCTGGTTTGCTAATTAATCTCCAGGAGCTATTTATTAGAACT 237
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 30

ADD26273

ID ADD26273 standard; DNA; 1134 BP.

XX AC ADD26273;

XX AC ADD26273;

DT 15-JAN-2004 (first entry)

XX Staphylococcus aureus dnaN DNA SEQ ID NO:78.

XX

Db 898 GAATGCTCTTACATCCAGAAATGGTACTGTAAAGAGAAAGTGTGATGCAACGAT 957
Qy 958 CAGTCTGGTAGTGAATTAACATATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTAAAA 1017
Db 958 GTTGAAGGTGGTAGCTGAAATTTCAITCAACTCTAAATATATATGATGATGCTTTAAAA 1017
Qy 1018 GCTATTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTGCGACCATTCACCTTA 1077
Db 1018 GCAATCGATATGATGAGGTTGAAGTTGAATTTCTTCGATACAAATGAAACCCATTATTCTA 1077
Qy 1078 ACACAGCGGATGAGGA 1094
Db 1078 AACCAAAAGGTGACGA 1094

RESULT 31
AAV74464
ID AAV74464 standard; DNA; 2347 BP.
XX AC
XX AAV74464;
XX 16-MAR-1999 (first entry)
XX Staphylococcus aureus contig SEQ ID #153.
XX
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
XX Staphylococcus aureus.
XX

Key Location/Qualifiers
FH 1741..1800
FT misc_feature
FT /*tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
XX
XX EP786519-A2.
XX
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97EP-00100117.
XX
XX 05-JAN-1996; 96US-0009861P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX
XX Claim 1; Page 775-776; 3271pp; English.
XX
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,

CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the computer
CC readable medium
XX
SQ Sequence 2347 BP; 865 A; 331 C; 382 G; 707 T; 0 U; 62 Other;
Query Match 20.8%; Score 236.2; DB 2; Length 2347;
Best Local Similarity 52.2%; Pred. No. 9.2e-41;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;
Qy 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAATCAACTAAACGT 60
Db 393 ATGATGGAATTCACCTATTAAAAAGAGATTATTTTATTACACAATTTAAATGACACATTAAAA 452
Qy 61 GCTATTAGCACTAAATGCGCATTCCTATTCTTTTCATCAATAAAAAATGAGTCACTTCT 120
Db 453 GCTATTTCACCAAGAACCAACATTACCTATATTAACCTGGTATCAAAATCGATCGGAAGAA 512
Qy 121 ACAGGAGTAACTTTAAACAGGGTCTAACGGTCAAAATATCAATTTGAAAAACACTATTCTCTGA 180
Db 513 CATGAAGTTATATTAACCTGGTTCAGACTCTGAAATTTCAATAGAAATCACATTCTCTAAA 572
Qy 181 AGT---AATGAAATGCTGGTTGCTAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCT 237
Db 573 ACTGTAGATGCGGAGATATTGTCAATATTTTCAAGAACAGGCTCAGTAGTACTTCTGGA 632
Qy 238 AGTTTATTTTAAATATTTTCAAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATTT 297
Db 633 CGATTCTTTGTTGATATTTAAAAAAATTTACCTGGTAAAGATGTTAAATTTATCTCAAAAT 692
Qy 298 GAAACAAACCAAGTTGTTTAAACAGGCTGCTAAATCAGAGATTACCTTTAAAAAGGAAAGAT 357
Db 693 GAACAAATTCAGACATTAATTTACATCAGCTCATCTCTGAATTTAAATTTAAGTGGCTTAGAT 752
Qy 358 GTTGACAGTATCTCTGCTCAAGAGATATCAAGAAATTCCTTTGATTTTAAAAACA 417
Db 753 CCAGATCAATATCTCTTATTACCTCAAGTTTCTAGAGATGACGCAATTTCAATTTGTCGTA 812
Qy 418 AAATTTTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTGT 477
Db 813 AAAGTGTCTTAAACAGGATGATGCAACAAACAAATTTTGCAGTGTCCACCTCAGAACACCC 872
Qy 478 CCTATTTTAAACAGGAGTTTCATATTGTTTAAAGTATATCAAAAGATTTTAAAGCAGTAGCG 537
Db 873 CCAGTACTAACTGGTGAACCTGGCTTATA---CAAGAAATGAAATTAATATGCACAGCG 929
Qy 538 ACTGACTCTCATCGTATGAGCCAAAGCTTTTAATCCTTTTGGACAATCTTTCAGCAGATTGG 597
Db 930 ACTGACTCACACCGCTTGGCTGTAAGAAAGTTGCAAGTTAGAGATGTTTCTGAAAAACAA 989
Qy 598 ATGTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATATT 657
Db 980 AATGTTCATCTCCAGGTAAGGCTTTAGCTGAATTTAAATAAAATTTATGTCTGACAAATGAA 1049
Qy 658 GAGACCGTTGAGGTATTTTCTCAACAAAGCAAAATCTTTGTTTCAAGAGTGAACACATTTCT 717
Db 1050 GAAGACATTGATATCTCTTCTTCAACCCAAAGTTTATTAAAGTTGGAATGTGAAC 1109
Qy 718 TTTTATACAGCCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTATATGACAGAA 777
Db 1110 TTTATTTCTCGATTATTAGAAGGACATTTATCTGTATACAAACAGCTTTTATTCCTCGAAAC 1169
Qy 778 TTTGAGACGGAGTTGTTTCAATACCAATCCCTTCCCGAGCTTATGGAAGTGCCTTC 837
Db 1170 TATGAAATTAATAAGTATGACAAATGGGAGTTTATCATGCGGATGATCGTGCCTCT 1229
Qy 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATT 897
Db 1230 TTATTAGCGGTGAAGGTGGTAAATTAACCGTTATTAAATTAAGTACAGGTGATGACGCTGTT 1289

QY 898 TCAGTCATGTTAACTCACCTGAGTGTGTAAGGTAAACGAGGATTTAGATATTGTTAGT 957
 Db 1290 GAATTTGCTCTTACATACACGAAATTTGTTACTGTAAAGAAAGTTGATGCAACGAT 1349
 QY 958 CAGTCTGTAGTATTAACATATCAGCTTCAATCCAACTTACCTTTATTGAGTCTTTAAAA 1017
 Db 1350 GTTGAAGGTGTAGCTGAAATTTCACTCACTAAATATATGATGATGCTTTAAAA 1409
 QY 1018 GCTATTAAAGTGAACAGTAAATTTCACTTCTATCACCAGTTCACCACTTACACCTTA 1077
 Db 1410 GCAATCGATATGATGAGTTGAAGTTGAATTTCTCGGTACAATGAACCAATTTATTCTA 1469
 QY 1078 ACACGAGCGGATGAGGA 1094
 Db 1470 AAACCAAAAGGTGACGA 1486

RESULT 32

AAZ31005
 ID AAZ31005 standard; DNA; 1134 BP.

AC AAZ31005;

XX 05-JAN-2000 (first entry)

DE Partial dnaN gene.

XX Gram positive bacteria; dnaE; dnaX; dnaB; dnaG; dnaN; dnaG; helicase;
 KW alpha subunit; DNA polymerase III holoenzyme; gamma subunit; tau subunit;
 KW clamp loader; glue protein; replication; antibiotic; ss.

OS Staphylococcus aureus.

Key Location/Qualifiers
 FT CDS 1..1134
 FT /*tag= a
 FT /note= "Dna N gene product"

XX WO9937661-A1.

XX 29-JUL-1999.

XX 25-JAN-1999; 99WO-US001547.

XX 27-JAN-1998; 98US-0074522P.

XX 22-JUL-1998; 98US-0093727P.

XX (UVRQ) UNIV ROCKEFELLER.

XX O'donnell ME, Zhang D, Whipple R;

XX WPI; 1999-590685/50.

XX P-PSDB; AAY49071.

XX New isolated dnaE, dnaX and dnaB genes from Gram positive bacteria, used
 to develop screening assays for identifying antibiotic compounds.

XX Example 8; Page 30; 132pp; English.

XX This sequence is the partial dnaN gene of Staphylococcus aureus. The
 CC invention relates to a number of isolated DNA molecules from Gram
 CC positive bacterium, corresponding to dnaE (AAZ31001), dnaX (AAZ31002), and
 CC dnaB (AAZ31003). The PolC, dnaN and dnaG genes (AAZ31004-231006) are also
 CC identified. The dnaE gene corresponds to the alpha subunit of the
 CC Escherichia coli, DNA polymerase III holoenzyme, dnaX corresponds to the
 CC gamma and tau subunits, and dnaB corresponds to the helicase. The alpha
 CC subunit is the actual DNA polymerase, the gamma complex forms the clamp
 CC loader and tau is a "glue protein". dnaX encodes both gamma and tau, Tau
 CC is the product of the full gene, while gamma is the product of the first
 CC two thirds of the gene. dnaN forms the beta subunit which forms the
 CC sliding clamp, and dnaG encodes a primase. The DNA sequences of the
 CC invention can be used to identify agents that inhibit or promote DNA
 CC replication by acting on various parts of the gram positive bacterial DNA

CC polymerase holoenzyme. The products and methods of the invention can be
 CC used for identifying pharmacological agents or lead compounds for agents
 CC active at the level of a replication protein function, particularly DNA
 CC replication. The agents identified can be used as antibiotics

XX Sequence 1134 BP; 411 A; 176 C; 194 G; 353 T; 0 U; 0 Other;
 SQ

Query Match 20.5%; Score 233; DB 2; Length 1134;
 Best Local Similarity 52.1%; Pred No. 3.9e-40;
 Matches 571; Conservative 0; Mismatches 520; Indels 6; Gaps 2;

QY 1 ATGATTCATTTTCAATTAATGCGACATTTATTTATTCATGCTTTTAAATFACAACTAAACGT 60
 Db 1 ATGATGGAATTCACATTTAAAGAGATTTATTTTATACACAATTAATGACACATTTAAA 60

QY 61 GCTATTAGCAGCTAAATGCGCATTCCTTCTTTCATCAATAAAATGAAGTCACCTTCT 120
 Db 61 GCTATTTCACCAAGAACCAACATTTACCTATATTAATGCTATCAAAATCGATGCGAAAGAA 120

QY 121 ACAGAGTAACCTTTAAACAGGCTTAACGGTCAAAATATCAATTTGAAACACATTTCTCTGTA 180
 Db 121 CATGAAGTTATATTAACCTGTTTCAGACTCTGAAATTTCAATAGAAATCCTATTTCTCTAAA 180

QY 181 AGT---AATGAAATGCTGTTGCTTAATTAACCTCTCCAGAGCTATTTTATTAGAAGCT 237
 Db 181 ACTGTAGATGCGAAGATATTTGTCATATTTTCAGAAACAGGCTCAGTAGTACTTCTCTGGA 240

QY 238 AGTTTTTTTATTAATTTATTTTCAAGTTTTCAGATATTTAGTATAAATGTTTAAAGAAAT 297
 Db 241 CGATTTCTTTGTTGATATTAATAAAATTAACCTGGTAAAGATTTAAATTTATCTACAAAT 300

QY 298 GAACAACACCAAGTTGTTTAAACAGTGGTAATTCAGATTTACCTTTAAAGGAAAGAT 357
 Db 301 GAACAATTCAGACATTAATTAATCATCAGTCTTCTGAATTTAATTTAGTGGCTTAGAT 360

QY 358 GTTGACCAAGTATCTCGTCTCAAGAAAGATATCAACAGAAATCTTTGATTTTAAAAACA 417
 Db 361 CCAGATCAATATCTTTTATTTACCTCAAGTTTCTAGAGATGACCAATTCATTTGTCGGTA 420

QY 418 AAATTTTGAAGTCTATTTATGCTGAAACAGCTTTTTCAGCCAGTTTCAAGAAAGATCGT 477
 Db 421 AAAGTACTTAAACACGTTGTTGCAAAACGAATTTTGCAGTGTCCACCTCAGAAACACGC 480

QY 478 CCTATTTTAAACAGGATTCATATTTGTTAAAGTAATCATTAAGATTTTAAACAGCAGTACGC 537
 Db 481 CCAGTACTAATCTGTTGTAAGTGGCTTATA---CAAGAAATGAATTAATATGACACAGC 537

QY 538 ACTGACTCTCATGTAAGCCAAACGTTTAAATCACTTTTGGACAATCTTTCAGCAGATTTG 597
 Db 538 ACTGATTCACACCGCTTGGCTGTAAGAAAGTTGCAAGTGTAGAGATGTTTCTGAAACAAA 597

QY 598 ATGTTAGTTCTTCCCAAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTT 657
 Db 598 AATGTCATCATTTCCAGGTAAAGCTTTAGCTGAATTTAAATTAATTTATGCTCACAATGAA 657

QY 658 GAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGAGTGAACACATTTCT 717
 Db 658 GAAGCATTTGATATCTTCTTCTTCAACCAAGTTTATTTTAAAGTTGAATGTGAAC 717

QY 718 TTTTATACAGCTCTTGAAGGAAATTTATCCGATACAGCCGTTTATTTAATAGAGAA 777
 Db 718 TTTATTTCTCGATTTATTAAGAGGCAATTTCTGTAACAACACAGTTTATTTCCCTGAAAC 777

QY 778 TTTGAGACGGAGTTGTTTTCATATCCCAATCCCTTCCGACGCTATCGAAGCTGCTTC 837
 Db 778 TATGAAATTAATTAAGTATAGACAATGGGAGTTTATCATGCAATGATGCTGCTCT 837

QY 838 TTGATTTCTAATGCTACTCAAAATGTTACTGTTAAGCTTGAGATTTACTCAAAATCATATT 897
 Db 838 TTATTAGCAGCGTGAAGGTGGTAAATTAACGTTTATTAATTAAGTACAGGTGATGACCTGTT 897

QY 898 TCAGCTCATGTTAACTCACCTGAGTGTGTAAGGTAAACGAGGATTTAGATATTGTTAGT 957

Db 898 GAATTATCTTCTACATCAGCAAAATTTGGTACTGTGTAAGAAAGAGTTGATGCAAAACGAT 957
Qy 958 CAGCTCGTGTAGTATTAACTATACAGCTTCAATCCAACTTACCTTATTGAGTCTTTAAAA 1017
Db 958 GTTGAAGGTGGTGGCTGAAATTTCAATCAACTTAAATATATGATGATGCTTTAAAA 1017
Qy 1018 GCTATTAAAGTGAACAGTAAATTCATTTCTTTATCACAGTTCGACCAATTCACCTTA 1077
Db 1018 GCAATCGATAATGATGAGGTGAAGTTGAATTTCTTCGGTACAAATGAAACCAATTTATTCTA 1077
Qy 1078 ACACGAGCGATGAGGA 1094
Db 1078 AAACCAAAAGGTGACGA 1094
RESULT 33
AAS54965
ID AAS54965 standard; DNA; 1134 BP.
XX AAS54965;
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #1277.
DE
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX
XX WO200170955-A2.
FN
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR P-PSDB; AAU37106.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 8602; 511bp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1134 BP; 414 A; 175 C; 192 G; 353 T; 0 U; 0 Other;
Query Match 20.5%; Score 233; DB 4; Length 1134;
Best Local Similarity 52.1%; Pred. No. 3.9e-40;
Matches 571; Conservative 0; Mismatches 520; Indels 6; Gaps 2;
Qy 1 ATGATTCAAATTTTCAATTAATGCGACATTAATTTATTTATTTATTTAAATACAACTAAACGT 60
Db 1 ATGATGGAAATTCACATTTAAAGAGATTTATTTATTTACAAATTAATTAATGACACATTAAAA 60
Qy 61 GCTATTAGCAGTCAAAAATGCCAATTCCTATTCTTTTCATCAATAAAAAATTTGAAAGTCACTTCT 120
Db 61 GCTATTTCCCAAGAACCAACATTAACCTATATACTGATTAATAAATCGATGCAAAAGAA 120
Qy 121 ACAGGAGTAATTTAAACAGGCTTAACGGTCAATATCAATTAATTAATAAAGTCACTTCTGTA 180
Db 121 CATGAAGTTATTAATTAACCTGTTTCAAGTCTGAAATTTTCAATAGAATACTATTTCTTAAA 180
Qy 181 AGT---AATGAAATGCTGGTTGCTAATTAACCTCTCCAGGAGCTATTTTATTTAGAACT 237
Db 181 ACTGTAGATGCGAGATATTGTCATATTTTCAGAAACAGGCTCAGTAGTACTTCTTCTGGA 240
Qy 238 AGTTTATTTTAAATATTAATTTCAAGTTTCCAGATATTAGTATAAATTTTAAAGAAATTT 297
Db 241 CGATTTCTTTGTGATATTATAAAAAAATTAACCTGTTAAAGATGTTAAATTTATCTACAAAT 300
Qy 298 GAACAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAGAT 357
Db 301 GAACAATTCACAGCATTAATTAACATCAGGTCATTTCTGAAATTTTAAATTTTAAAGTGGCTTAGAT 360
Qy 358 GTTGACAGATATCTCGTCTACAAGAGTATCAACAGAAAAATCCTTTGATTTTAAAAACA 417
Db 361 CCAGATCAATATCTTTTATTAACCTCAAGTTTCTAGAGATGAGCAATTTCAATTTGTCGTA 420
Qy 418 AAATTTATTTGAAGTCTATTATTGCTGAAACAGCTTTTTCAGCCAGTTTACAGAAAGTGGT 477
Db 421 AAAGTGTCTTAAAAACGTAATTTGCAAAACGAAATTTTGCAGTGTCCACCTCAGAAAACAGC 480
Qy 478 CCTATTTTAAACAGGATTCATATTGTATTAGTAAATCAATAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTAACTGGTGAACCTGGCTTATA---CAAAGAAATGAATTAATATGCAACGCG 537
Qy 538 ACTGACTCTCATCGTATGAGCAACGTTTAACTCTTGGCAATATCTTTGAGCAATTTG 597
Db 538 ACTGACTCACACCGCTTGGCTGTAAGAAAGTTGCAGTTAGAGATGTTTCTGAAAAACAA 597
Qy 598 ATGGTAGTTCTTCCAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATATT 657
Db 598 AATGTATCATTTCCAGGTAAAGCTTTAGCTGAAATTAATAAATAAATTTATGCTTGACAATGAA 657
Qy 658 GAGACCGTTGAGTATTTTCTCACCAGCAATCTTTGTTGAGAGTGAACACATTTCT 717
Db 658 GAAGACATTTGATATCTTCTTTGCTTCAACCAAGTTTATTTTAAAGTTGGAATGTGAAC 717
Qy 718 TTTTATACACCGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
Db 718 TTTTATTTCTCAATTTATAGAAGGACATTTCTGATACACACGTTTATTTCTGAAAAAC 777
Qy 778 TTTGAGACGGAGTTGTTTTCATACCCCAATCCCTTCGCCACGCTATGGAACGTCCTTC 837
Db 778 TATGAAATTAATAAGTATAGACAAATGGGAGTTTATCATGCGATTTGATCGTGCATCT 837
Qy 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTCAAAATCATATT 897
Db 838 TTATTTAGCAGTGAAGGTGGTAAATCAACGTTATTAATTAATTAAGTACAGGTGATGACGTTGT 897
Qy 898 TCAGTCTATGTTAACTCACTACCTGAGTTGGTGAAGTAAACAGGAGTTTATAGATTTTGTAGT 957
Db 898 GAAATATCTTCTACATCACCAAGAAATTTGGTACTGTAAAAGAAAGTTGATGCAACGAT 957

DB 1078 AACCAAAAGGTGACGA 1094

RESULT 35

ACA20030

ID ACA20030 standard; DNA; 1134 BP.

XX ACA20030;

AC ACA20030;

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #1687.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

KW Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall L, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-029926/02.

DR P-PSDB; ABU16160.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 7900; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1134 BP; 411 A; 178 C; 194 G; 351 T; 0 U; 0 Other;
Query Match 20.5%; Score 233; DB 8; Length 1134;
Best Local Similarity 52.1%; Pred. No. 3.9e-40;
Matches 571; Conservative 0; Mismatches 520; Indels 6; Gaps 2;
QY 1 ATGATTCAATTTTCAATTAATCGCACATTATTATTATTCATCTGCTTTAAATACAACATAACCGT 60
DB 1 ATGATGGAATTCACATTTTAAAGAGATATTATTATACACATTTAAATGACACATTAAAA 60
QY 61 GCTATTAGCATTAAAAATGCCATTCTCTATTCTTTTCAATCAATAAAAAATGGAAGTCATCTT 120
DB 61 GCTATTTCACCAAGAACCAACATTACCTATATTAACTGGTATCAAAATCGATCGGAAGAA 120
QY 121 ACAGAGTAACCTTTAAACAGGGTCTAACGGTCAATATCAATTTGAAACACATTTCTCTGTA 180
DB 121 CATGAAGTTATATCTAACTGGTTACAGCTCTGAAATTTCAATAGAAATCACTATTCTAAA 180
QY 181 AGT---AATGAAATGCTGGTTTGCTAATTACCTCTCCAGGAGCTATTATTTATAGAGCT 237
DB 181 ACTGTAGATGGGAAGATATTGTCATATTTCAGAAACAGGCTCAGTAGTACTTCTTGA 240
QY 238 AGTTTTTTTATTAATATTATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAAT 297
DB 241 CGATTCTTTGTTGATATTATAAAAAAATTAACCTGGTAAAGATGTTTAAATTTATCTACAAAT 300
QY 298 GAACAAACCAAGTTGTTTTAACCAAGTGGTAAATCAGAGATTACCTTAAAGGAAAGAT 357
DB 301 GAACAAATTCACAGACATTAATTAATCACTCAGGTCATTTCTGAATTTTAAAGTGGCTTAGAT 360
QY 358 GTTGACCAAGTATCCTCGTCTACAAGAGTATCAACAGAAATTCCTTTGATTTTAAAAACA 417
DB 361 CCAGATCAATATCCTTTTATTACCTCAAGTTCTAGAGATGCGCAATTCATTTGTCGTA 420
QY 418 AAATTTATGAAAGTCTATTATTGCTGAAACAGCTTTTGCGCCAGTTTACAAGAAAGTCGT 477
DB 421 AAAGTGCTTAAAAACGTTGTCACAAACGAATTTTGGCAGTGTCCACCTCCAGAAACACGC 480
QY 478 CCTATTTTAACAGGAGTTTCATATTGTTATTAAGTATCATTAAGATTTTAAAGCAGTAGCG 537
DB 481 CCAGTACTTAACCTGGTGTGAATGCTGCTTATA---CAAGAAATGAAATTAATATGCAACGG 537
QY 538 ACTGACTCTCATCGTATGAGCCAAAGTTTAACTCACTTTGGCAATACCTTCACAGATTTG 597
DB 538 ACTGACTCACACCGCTTGCTGTAAGAAAGTTGCAAGTTAGAAGATGTTTCTGAAACAAA 597
QY 598 ATGGTAGTTCTTCCAAAGTAAATCTTTTGAGAAATTTTCAGCAGATATTACAGATGATTT 657
DB 598 AATGTCATCAATCCAGGTAAGGCTTTAGCTGAATTAATAAATAAATTTATGCTGCAATGAA 657
QY 658 GAGACCGTTGAGTATTTTCTCACCAGCCAAATCTTTGTTGAGAGTGAACACATTTTCT 717
DB 658 GAAGACATTTGATATCTCTTTGCTTCAAAACCAAGTTTATTTAAAGTTGGAATGTGAAC 717
QY 718 TTTTATACACGCTCTTGAAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
DB 718 TTTTATTTCTGATTTATAGAGGACATTTATCTGATACACACGTTTATTCCTGAAAC 777
QY 778 TTTGAGACGGAGTTGTTTTCAATACCAATCCCTTCGCCACGCTTATGGAACGTCCTTC 837
DB 778 TATGAAATTAATAAGTATAGACAATGGGAGTTTATCATGCGATTGATCGTGCCTCT 837
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAGCTTTGAGATTTACTCAAAATCATATT 897
DB 838 TTTATAGACCGTGAAGGGTGAATAACCGTTATTAATTAATTAAGTACAGGTGATGACGTTGTT 897
QY 898 TCAGCTCATGTTAACTCACTCAGCTGAGTTCGTTAGGTAAACAGGAGATTTAGATATTGTTAGT 957
DB 898 GAAATTTCTTACATCACCAAGAAATTTGTTACTGTAAAAGAGAGTTGATGATCAACAGAT 957

QY 958 CAGTCTGTGAGTGAATTAATCACTACAGCTTCAATCAACTTACCTTATGAGCTTTTAAAA 1017
 Db 958 GTTGAAGTGGTGGCTGAAATTTCAATCACTTAAATATATGATGAGCTTTTAAAA 1017
 QY 1018 GCTATTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTGACCAATTCACCTTA 1077
 Db 1018 GCAATCGAATAATGATGAGGTTGAAGTTGAATTCCTCGGTACAATGAACCAATTTATCTA 1077
 QY 1078 ACACGAGCGATGAGGA 1094
 Db 1078 AAACCAAAAGGTGACGA 1094

RESULT 36
 AAD62920

ID AAD62920 standard; DNA; 1134 BP.

XX AC AAD62920;

DT 12-FEB-2004 (first entry)

Staphylococcus aureus dnaN gene beta subunit DNA.

KW Polymerase III enzyme; dnaE; dnaX; dnaB; Gram positive bacteria;
 KW drug discovery; antibiotic activity; gene; ds.

OS Staphylococcus aureus.

FH Key Location/Qualifiers

FT CDS 1. 1134

FT /*tag= a

FT /product= "Staphylococcus aureus dnaN beta subunit"

PN US2003129633-A1.

XX 10-JUL-2003.

XX 28-OCT-2002; 2002US-00282287.

XX 13-FEB-1998; 98US-0074572P.

XX 22-JUL-1998; 98US-0093727P.

XX 22-JAN-1999; 99US-00235245.

XX (ODON/) O'DONNELL M E.

PA (ZHAN/) ZHANG D.

PA (WHIP/) WHIPPLE R.

XX O'donnell ME, Zhang D, Whipple R;

PI WPI; 2003-829557/77.

XX P-PSDB; ABW01648.

XX New DNA replication proteins (i.e. subunits of the Staphylococcus aureus
 PT DNA polymerase III enzyme) and genes, useful in drug discovery to screen
 PT large libraries of chemicals for identification of compounds with
 PT antibiotic activity.

XX Disclosure; Page 17-18; 69pp; English.

XX The invention relates to an isolated polypeptide, which comprises at
 CC least one functionally active subunit of a Staphylococcus aureus DNA
 CC polymerase III enzyme. The subunit comprises a 573 residue dnaE amino
 CC acid sequence, a 566 residue dnaX amino acid sequence and/or a 457
 CC residue dnaB amino acid sequence. The proteins and nucleic acids
 CC replicate the chromosome of Gram positive bacteria and are useful in drug
 CC discovery to screen large libraries of chemicals for identification of
 CC compounds with antibiotic activity. The present sequence is S. aureus
 CC dnaN gene beta subunit DNA

XX Sequence 1134 BP; 411 A; 176 C; 194 G; 353 T; 0 U; 0 Other;

Query Match 20.5%; Score 233; DB 10; Length 1134;

Best Local Similarity 52.1%; Pred. No. 3.9e-40;

	Matches	571;	Conservative	0;	Mismatches	520;	Indels	6;	Gaps	2;
QY	1	ATGATTCAATTTTCAATTAATCGACATTTATTTATTTATGCTTTTAAATACAACTAAACGT	60							
Db	1	ATGATGGAATTCACATTTAAAGAGATTTATTTATTTATACAAATTAATGACACATTTAAA	60							
QY	61	GCTATTAGCAGTAAATGCGCATTCCTATTTCTTTCATCAATAAAATTTGAAGTCACTTCT	120							
Db	61	GCTATTTTCCACCAAGAACCAATACCTATATTAATGCTGATCAAAATCGATCGGAAGAA	120							
QY	121	ACAGGAGTAACTTTAAACAGGGTCTAACCGTCAATATCAATTTGAAAAACACTTCTCTGA	180							
Db	121	CATGAAGTTATATTAATTAAGTTCAGACTCTGAAATTTCAATAGAAAACTACTTATCTTAA	180							
QY	181	AGT---AATGAAAAATGCTGTTTCTAATTAACCTCTCCAGGAGCTATTTTATTAAGAAGCT	237							
Db	181	ACTGTAGATGGGAGATATTGTCAATATTTTCAGAAACAGGCTCAGTAGTACTTCTCTGGA	240							
QY	238	AGTTTTTTTATTAATATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAAT	297							
Db	241	CGATTTCTTTGTTGATTTATATAAAAAATTTACCTGTTAAAGATGTTAATTTATCTACAA	300							
QY	298	GAAACAACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTTAAAGGAAAAGAT	357							
Db	301	GAACAATTCAGACATTAATTTACATCAGGTCAITTTCTGAATTTAATTTGAGTGGCTTAG	360							
QY	358	GTTGACAGTATCTCTCGTCTCAAGAAAGTATCAACAGAAAAATCTTTGATTTTAAAAACA	417							
Db	361	CCAGATCAATATCTTTTATTAACCTCAAGTTTCTAGAGATGACGCAATTTCAATTTGTCG	420							
QY	418	AAATTTTGAAGTCTATTTATTTGCTGAAACAGCTTTTTCGACCCAGTTTACAGAAAGTC	477							
Db	421	AAAGTACTTAAACAGTGATTTGCACAAACGAATTTTTCAGTGTCCACCTCAGAAACACG	480							
QY	478	CCTATTTTAAACAGAGTTCAATTTGTTAATTAAGTAATCATAAAGATTTTAAAGCAGTAG	537							
Db	481	CCAGTACTAACTGCTGTGAATGCTGCTATA--CAAGAAATGAAATTAATGTCACAGCG	537							
QY	538	ACTGACTCTCATGATAGCCCAAGTTTAAATCACTTTTGACAAATCTTTCAGCAGATTTG	597							
Db	538	ACTGATTCACACCGCTTGGCTGTAAAGAGTTGACGTAGAGATGTTTCTGAAAAACAA	597							
QY	598	ATGCTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGAT	657							
Db	598	AATGTCATCATTCAGGTAAAGCTTTAGCTGAATTAATAAATTAATGCTGACAAATGAA	657							
QY	658	GAGACGCTTGAGGTATTTTCTCACAAAGCCAAATCTTTCAGAGAGTGAACACATTTCT	717							
Db	658	GAAAGACATTTGATATCTTCTTCTTCAAAACCAAGTTTATTTTAAAGTTGGAATGTGA	717							
QY	718	TTTTATACACGCTCTTAGAAGGAATTTATCCGATACAGACGGTTTATTAATGACAGAA	777							
Db	718	TTTTATTTCTGATTTAGAGGCAATTTCTGATACAAACAGTTTATTTCTCCGAAAC	777							
QY	778	TTTGAGACGGAGGTGTTTTTCAATCCCAATCCCTTCGCCACGCTATGGAAGTGCCTTC	837							
Db	778	TATGAAATTTAAATTAAGTATAGCAATGGGAGTTTATCATGCGATGATCGTCTCT	837							
QY	838	TTGATTTCTTAATGCTACTCAAAATGCTGTTTAAAGTTGAGATTACTCAAAATCATATT	897							
Db	838	TTATTAGACGTTGAAGGTGTTAATTAACGTTTATTAATTAAGTACAGGTGATGACGTTGT	897							
QY	898	TCAGCTCATGTTTAACTCACCTGAGGTGTTGTAAGGTAAACGAGGATTTAGATATTGTTAG	957							
Db	898	GAAATATCTTCTACATCACCAGAAATGCTGTTAAAGAGAGAGTTGATGTCMAACGAT	957							
QY	958	CAGTCTGTGAGTGAATTTAACTATCAGCTTCAATCAACTTACCTTTATGAGTCTTTAAAA	1017							
Db	958	GTTGAAGGTGGTGGCTGAAATTTTCAATCAACTTAAATATATGATGATGATGCTTTAAA	1017							
QY	1018	GCTATTTAAAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCAATTCACCTA	1077							
Db	1018	GCAATCGATATGATGAGGTTGAAGTTGAATTTCTTCGGTACATGAACCAATTTATTTCTA	1077							


```

Db      1078 CCAAAAGTGACGA 1091
|||||
RESULT 38
ABQ70950
ID      ABQ70950 standard; DNA; 4736 BP.
XX
AC      ABQ70950;
XX
DT      29-AUG-2003 (revised)
DT      29-AUG-2002 (first entry)
XX
DE      Listeria monocytogenes 4b contig DNA sequence #892.
XX
KW      Antibacterial; Listeria; food contamination; mutational analysis;
KW      infection; ds.
XX
OS      Listeria monocytogenes ATCC 19115.
XX
PN      WO200228891-A2.
XX
PD      11-APR-2002.
XX
PF      04-OCT-2001; 2001WO-FR003061.
XX
PR      04-OCT-2000; 2000FR-00012697.
XX
PA      (INSP ) INST PASTEUR.
PA      (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI      Kunst F, Glaser P;
XX
DR      WPI; 2002-332479/37.
XX
PT      New genomic sequences from Listeria species, useful for detection,
PT      treatment and prevention of infection, also related polypeptides,
PT      antibodies and modulators.
XX
PS      Claim 14; SEQ ID NO 3763; 180pp; French.
XX
CC      The present invention relates to nucleic acid sequences (ABQ67198-
CC      ABQ71212) from Listeria sp. The sequences are useful as probes and
CC      primers for identification and/or detection of Listeria (e.g. as
CC      contaminants in foods, or mutational analysis) and for analysis of gene
CC      expression. Proteins encoded by the nucleic acid sequences can be used to
CC      screen for compounds that modulate gene expression, replication and
CC      pathogenicity of Listeria (potential therapeutic agents), also for
CC      treating infections by Listeria, and are useful as immunogens in anti-
CC      Listeria vaccines. Note: The sequence data for this patent did not form
CC      part of the printed specification, but was obtained in electronic format
CC      directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC      on 29-AUG-2003 to standardise OS field)
XX
SQ      Sequence 4736 BP; 1521 A; 800 C; 938 G; 1477 T; 0 U; 0 Other;

Query Match      17.7%; Score 200.4; DB 6; Length 4736;
Best Local Similarity 51.5%; Pred No. 4.5e-33;
Matches 565; Conservative 0; Mismatches 521; Indels 12; Gaps 4;

QY      2 TCATTCAATTTTCAATTAATCGCACATTTATTTCATGCTTTTAAATACAACTAAACGCTG 61
Db      1218 TCATGAATTTTGTATTATGAGCGTGATCGTCTGTGCAAGCAGTCAATGAAGTACTCGTG 1277
QY      62 CTAATTAGCACATAAAATCCCAATTCCTATCTTTTCATCAATAAATGAAGTCACTCTA 121
Db      1278 CCATCTCTGCAAGACACAGCAATCAATTTTAAACGGGATAAATAAGTCTGAATGATG 1337
QY      122 CAGAGTAATCTTAAACAGGGTCTAACGGTCAAAATATCAATTTGAAACACATCTTCCGTGA 181
Db      1338 AAGGTGTAACACTAACTAGGTAGTATTCGATATTTCAATCGAAGCAATTTATTCATTAA 1397
QY      182 GT---AATGAAATGCTGGTTTGTCTAATTACCTCTCCAGGAGCTATTTTATTAGAGCTA 238

```

```

Db      1398 TTGAAAATCATGAGTAATTTGTAGAGTGGAGAGTTTTGGTGAATTTGTACTTCAATCAA 1457
QY      239 GTTTTTTTTAAATTAATATTTCAGAGTTTGGCCAGATTAATTAGTATAAATGTTAAAGAAATG 298
Db      1458 AATACCTTTGGCGATATTGTTCTGTTTACCAGAGAAATATGTAAGAAATGAAGTGACTT 1517
QY      299 AACAAACACCAAGTTGTTTTAACCACTGTTAAATCAGAGATTACCTTTAAAGGAAAAGATG 358
Db      1518 CTAACCTACCAACCAACATTAAGTTCTGGCCAGCAATCTTTTACATTAATGGCTTAGATC 1577
QY      359 TTGACCAAGTATCTCTCTACAAAGAGTATCAACAGAAAATCCCTTTGATTTTTAAACAA 418
Db      1578 CAATGGAATATCTTAAATTAATCTGAAGTAAACAGACGGAACAAATTAATTTCAATTA 1637
QY      419 AATTAATCAAGTCTAATTAATTTGCTGAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCGTC 478
Db      1638 ATGTACTTAAATAATTTGTTAGACAAATCTGTTTGTCTGCTGCGATTGAAGTTCGTC 1697
QY      479 CTATTTTAAACAGGAGTTTCATATTGTTAATTAATCAATAAGATTTTAAAGCAGTAGCGA 538
Db      1698 CAGTACTTACTGTTGTAACCTGGATTAACAAGAAAATAAA---CTAAGCGCAGTTGCAA 1754
QY      539 CTGACTCTCATCTATGAGCCACGTTTAACTTTGAGACA---ATACTTTCAGCAGATT 595
Db      1755 CCGATAGTCATCGTCTAGCTTTACGTGAAATACCTCTTGAAACAGACATTTGATGAAGAT 1814
QY      596 TGATGGTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATA 655
Db      1815 ACAACATTTGTTTCTCTGGAATAAGTTTATCTGAATTAATAAATCTTTAGATGATGCAA 1874
QY      656 TTGAGACGGTTGAGGTATTTTCTCACAAGCCAAATCTTTGTTTCAGAAGTGAACACATTT 715
Db      1875 GCGAATCTATTGAAATGACCCCTTGCCAAACCAACCAATCTTTTAAATTAAGATTTAT 1934
QY      716 CTTTATTATACAGCCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAG 775
Db      1935 TATTTATTCTCGTTTACTTGAAGGTAGTTATCCAGATACATCTCGATTAAATTTCCAACGTG 1994
QY      776 AATTGAGACGGAGTTGTTTTCAATACCCAAATCCCTTCGCCACGCTATGGAAACGTGCT 835
Db      1995 ATACTAAATCAGAAATGATGATTAATCCAAAGCAATTTTACAGCAATGACCGTGCCT 2054
QY      836 TCTTGATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTAATCTCAAAAT---C 892
Db      2055 CCTACTTGTCTCGGAAATTCGTAATGATTAATTAATTAATGACGCTAGAAAATGGCC 2114
QY      893 ATATTTCACTCATGTTTAACTCACCTGAGTTGGTAAAGTAAACGAGGATTTAGATATTG 952
Db      2115 AAGTAGAAGTATCATCCAAATCTCCAGAGTTGGGAAATGTTTCTGAAAATGCTTCAGCC 2174
QY      953 TTAGTCAGTCTGGTAGTATTAATCAATCAGCTTCAATCAACTTACCTTATTTAGTCTTT 1012
Db      2175 AAGTTTTTACTGGCGAAGAGATCAAAATATCTTTTAAACGGTAAATATACATGATGATGCC 2234
QY      1013 TAAAGCTATTAAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTCGACCAATCA 1072
Db      2235 TAGGTGCTTTTGAAGTGATGATTAATCAAAATTCCTTCTCCGTACAAATGACCAATTCG 2294
QY      1073 CCCTAACACCGCGGATG 1090
Db      2295 TACTTCGACCAAAAGATG 2312

```

```

RESULT 39
ABQ67196.2/c
Continuation (3 of 7) of ABQ67196 from base 200001 (Listeria innocua contig DNA sequence
WP Sequence split into 7 fragments LOCUS ABQ67196 Accession ABQ67196
WP Fragment Name      Begin      End
WP ABQ67196.0          1      110000
WP ABQ67196.1        100001    210000
WP ABQ67196.2        200001    310000
WP ABQ67196.3        300001    410000

```

```

WP ABQ67196_4 40001 510000
WP ABQ67196_5 50001 610000
WP ABQ67196_6 60001 684707

Query Match
  17.5%; Score 198.8; DB 6; Length 110000;
Best Local Similarity 51.4%; Pred No. 1.7e-32;
Matches 564; Conservative 0; Mismatches 522; Indels 12; Gaps 4;

Qy 2 TGATTCAATTTCAATTAATCGCATATTATTTATTCATGCTTTAAATAACAATAAAAGT 61
Db 67740 TCATGAAATTTGTTATTGAGCGTGATCGTCTGTCCAGAGCATCAATGAAGTTACTCGT 67681

Qy 62 CTATTAGACATAAATAATCCATCTCTATTCTTCATCAATAAATAATGAAGTCACTTCTA 121
Db 67680 CCATCTCTGCAAGAACCAAGATTCATTTTAAACGGGATGATAAATAGCTGTAATGATG 67621

Qy 122 CAGGAGTAATTTTAAACAGGCTTAAACGGTCAATCAATCAATTAATGAATACTATCTCTGTA 181
Db 67620 AAGGTGTTTAAACCGGTAGCGATTCGTATATTTCCATCGAAGCATTTATTCATTA 67561

Qy 182 ---GTAATGAAATGCTGGTTTGCTTAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTA 238
Db 67560 TCGAAATGATGAAGTAATTTGAGAAGTTGAAAGTTTGGTGGTATCGTCTCTCAATCCA 67501

Qy 239 GTTTTATTTAATATTAATTTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATTG 298
Db 67500 AGTACTTCGGTGATATGTTCTCGTGTTCCTGCTGAGAAATGTCGAAATTTGAAGTTACTA 67441

Qy 299 AACAAACCAAGTTGTTTTTAAACAGTGTGTAATCAGAGATTACCTTTAAAGGAAAGATG 358
Db 67440 CAATTTCAACTCAACTAATTAATGTTCTGCCAAGCTTCTTTACACTAAACGGCTTAGATC 67381

Qy 359 TTGACAGTATCTCTGCTACAAGAGATCAACAGAAATTCCTTTGATTTTAAACAA 418
Db 67380 CAATGGAATACCTAAATACCAGAGATTAAGTGGTAAATAATTAATTAATTCCTATTA 67321

Qy 419 AATATTGAAGTCTATTATGCTGAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCGTC 478
Db 67320 ATGCTTTTAAACAACTAATTAAGCAAACTGTTTGTGCTATCGAAGTTGCC 67261

Qy 479 CTATTTTAAACAGGAGTTCATATTTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGA 538
Db 67260 CAGTCTTACAGGTGTTAAGTGGATTAAT---TAAAGATTAACAACCTTAGCGAGTTGCA 67204

Qy 539 CTGACTCTCATGATGACCAACGATTTAATCACTTTGGA---CAATCTTACAGCATTT 595
Db 67203 CCGATGATCATGCTGATGATTTACGTGAAATTCACCTTTGAAACAAATATTGACGAAGAT 67144

Qy 596 TGATGGTATGTTCTTCAAGTAATCTTTGAGAGAAATTTTACAGATATTACAGATGATA 655
Db 67143 ACAACATTTGTTATTCAGGAAAGCTTTAGCTGAACTAAATAAATTTTACAGCATGCA 67084

Qy 656 TTGAGACCGTTGAGGTATTTTCTCACCAGGCAATCTTGTTCAGAAGTGAACACATTT 715
Db 67083 GTGAATCAATGAATGACACTCGCAATTAATCAATCTTATTTAAATTAAGATTTAT 67024

Qy 716 CTTTATTATACGCTCTTTAGAGGAAATTTCCCGATACAGACCGTTTATTAATGACAG 775
Db 67023 TATTTTATCTGTTTACTGGAAGGTAGTTACCCAGATACATCAAGCTTAATTTCCAACTG 66964

Qy 776 AATTTGAGACGGAGTTGTTTTCATACCCATCTTCCGACGCTATGAAAGTGCCT 835
Db 66963 ATACTAAATCAGAACTTGTCTAATTAATCCAGAGCAATTTTACAGGATTTGACCGTGT 66904

Qy 836 TCTTGATTTCTAATCTACTCAATGTTGTTAAGCTTGAGATTTCAAAAT---C 892
Db 66903 CCTACTTGTCTCGGAAACCGTAATTAAGTTTAAATTAATGACCTTGAAATGAC 66844

Qy 893 ATATTTCAGCTCATGTTAACTCACCTGAGGTGGTGAAGTAACGAGGATTTAGATTTG 952
Db 66843 AAGTTGAAGTATCCTCTAATCTCTGCTGAGTTGGGAATGTTTCGGAAATATGCTTTTAGTC 66784

Qy 953 TTAGTCAGTCTGGTAGTGAATTAACATATACAGCTTCAATCCAACTTACCTTATTGAGTCTT 1012

```

```

Db 66783 AAAGCTTTACTGGTGAAGAAATTAATAATATCCTTTAAACGGTAAATACATGATGTCAT 66724
Qy 1013 TAAAGCTATTAAAGTGAACAGTAAATAATTTCTTTATCACCAGTTCCGACCATTC 1072
Db 66723 TGCGCGCATTTGAAGGCGATGATATTCAAATCTCTCTCTGGCACAATGAGACCTTTCG 66664
Qy 1073 CCCTAACACACGCGCATG 1090
Db 66663 TACTTCGACCAAAAGATG 66646

RESULT 40
ABQ69245_00
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245
WP Fragment Name Begin End
WP ABQ69245_00 1 110000
WP ABQ69245_01 10001 210000
WP ABQ69245_02 20001 310000
WP ABQ69245_03 30001 410000
WP ABQ69245_04 40001 510000
WP ABQ69245_05 50001 610000
WP ABQ69245_06 60001 710000
WP ABQ69245_07 70001 810000
WP ABQ69245_08 80001 910000
WP ABQ69245_09 90001 1010000
WP ABQ69245_10 100001 1110000
WP ABQ69245_11 110001 1210000
WP ABQ69245_12 120001 1310000
WP ABQ69245_13 130001 1410000
WP ABQ69245_14 140001 1510000
WP ABQ69245_15 150001 1610000
WP ABQ69245_16 160001 1710000
WP ABQ69245_17 170001 1810000
WP ABQ69245_18 180001 1910000
WP ABQ69245_19 190001 2010000
WP ABQ69245_20 200001 2110000
WP ABQ69245_21 210001 2210000
WP ABQ69245_22 220001 2310000
WP ABQ69245_23 230001 2410000
WP ABQ69245_24 240001 2510000
WP ABQ69245_25 250001 2610000
WP ABQ69245_26 260001 2710000
WP ABQ69245_27 270001 2810000
WP ABQ69245_28 280001 2910000
WP ABQ69245_29 290001 3010000
WP ABQ69245_30 300001 311208
ID ABQ69245 standard; DNA; 3011208 BP.
XX AC ABQ69245;
XX DT 29-AUG-2002 (first entry)
XX DE Listeria innocua DNA sequence #684.
XX KW Antibacterial; Listeria; food contamination; mutational analysis;
XX OS infection; ds.
XX PN Listeria innocua.
XX WO WO200228891-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-FR003061.
XX PR 04-OCT-2000; 2000FR-00012697.
XX PA (INSP ) INST PASTEUR.
XX PI (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Kunst F, Glaser P;
XX

```

DR WPI: 2002-332479/37.
 XX New genomic sequences from *Listeria* species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators.
 XX
 PS Claim 5; SEQ ID NO 2058; 180pp; French.
 XX
 CC The present invention relates to nucleic acid sequences (AB067188-
 CC AB071212) from *Listeria* sp. The sequences are useful as probes and
 CC primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of gene
 CC expression. Proteins encoded by the nucleic acid sequences can be used to
 CC screen for compounds that modulate gene expression, replication and
 CC pathogenicity of *Listeria* (potential therapeutic agents), also for
 CC treating infections by *Listeria*, and are useful as immunogens in anti-
 CC *Listeria* vaccines. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3011208 BP; 941651A; 568176C; 559189G; 942192T; 0U; 0Other;
 Query Match 17.5%; Score 198.8; DB 6; Length 110000;
 Best Local Similarity 51.4%; Pred. No. 1.7e-32;
 Matches 564; Conservative 0; Mismatches 522; Indels 12; Gaps 4;
 QY 2 TGATTCAATTTTCATTAATCGCACATATTATTATTCATGCTTTAAATACAACTAAACGTG 61
 Db 1866 TCATGAATTTGTTATTGAGCGTGATCGTCTGTGTCGAAGCAGTCAATCAAGTTACTCGTG 1925
 QY 62 CTATTAGCACTAAAATGCCATTCCTATTCTTTCATCAATAAAAATGAACTCACTTCTA 121
 Db 1926 CCATCTCTGAAGAACAAACGATTCCAATTTACGGGNTAAATATAGTCGTAAATGATG 1985
 QY 122 CAGAGTAACCTTTAAACAGGCTTAACCGTCAAAATATCAATTGAAAAACACTATTCTCTGTA 181
 Db 1986 AAGGTGTACTTTAAACCGGTGAGCGATTCTGATATTTCCATCGAAGCATTTATTCCATTAA 2045
 QY 182 ---GTAATGAAATGCTGGTTGCTAATACCTCTCCAGCAGCTATTTTATTAGAACGTA 238
 Db 2046 TCGAAATGATGAAGTAATTTGTAGAGTTGAAAGTTTTGGTGGTATCGTTCTTCAATCCA 2105
 QY 239 GTTTTTTTAATATATATTTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTG 298
 Db 2106 AGTACTTCGGTGATATGTTTCGTTCGTCTCAAGAAATGTCGAAATTTGAAGTTTACTA 2165
 QY 299 AACACACCAAGTTGTTTTAAACCAAGTGTTAAATCAGAGATACAGATATACCTTTAAAGAAACAA 418
 Db 2166 CAAATTTACAACTAACATTAGTTCTGGCAAGCTTCCTTTACACTAAACGGCTTAGATC 2225
 QY 359 TTGACCAAGTATCTCTGCTTACAAGAGATATCAACAGAAATCCCTTTGATTTTAAACACAA 418
 Db 2226 CAATGGAATACCCATAAATACCAGAAAGTTACTGATGGTAAAAATATTTAAATTCCTATT 2285
 QY 419 AATTTATTGAAGTCTATTATTGCTGAAACAGCTTTTCGCCAGCTTTTACAAGAAAGTCGTC 478
 Db 2296 ATGTCCTTAAAAACATTATAGCAAACTGTTTTCCTGTTCTGCTATCGAAGTTTCGCC 2345
 QY 479 CTATTTTAAACAGAGTTCTATATTGTTAAGTAATCATAAAGATTTTAAAGCAGTAGCGCA 538
 Db 2346 CAGTGCCTTACAGGTGTTAACTCGAATAT --TAAAGATAACAACTTAGCGCAGTTGCCA 2402
 QY 539 CTGACTCTCATGATGAGCCAAAGCTTTTAACTCACTTTGGA ---CAATCTTCAGCAGATT 595
 Db 2403 CCGATTAGTCAGCTCTAGCAATACGTTAGGAAATTCACCTTGAAACAAATATTGACGAAGAT 2462
 QY 596 TGATGGTATGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATA 655
 Db 2463 ACAACATTTGTTATTCAGGAAAAAGCTTAGCTGAACATAAATAATTTAGACGATGCCA 2522
 QY 656 TTGAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCCAGAGTGAACATTT 715
 Db 2523 GTGAATCATGAATGAACACTCGCAAAATAATCAATCTTATTTAAATTTAAAGATTTTAT 2582

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 10:52:07 ; Search time 3366 Seconds
(without alignments)
15931.824 Million cell updates/sec

Title: US-10-048-071-27
Perfect score: 1134
Sequence: 1 atgattcaatttcattaa.....ttacaccagtagcaaac 1134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sv.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1134	100.0	1134	1	AF280765 Streptoco
2	1126	99.3	1134	6	CQ648158 Sequence
3	1126	99.3	12551	1	AE006472 Streptoco
4	1118	98.6	110000	1	CP000003 Streptoco
5	1116.4	98.4	14943	1	AE009953 Streptoco
6	1114.8	98.3	51896	1	AE014136 Streptoco
7	1114.8	98.3	311600	1	AP005141 Streptoco
8	849	74.9	3200	6	AR157860 Sequence
9	849	74.9	3200	6	E17380 gDNA derive
10	690.8	60.9	14925	1	AE014853 Streptoco
11	641.2	56.5	1134	6	CQ648156 Sequence
12	641.2	56.5	1137	6	AX606582 Sequence
13	641.2	56.5	21362	1	AE014191 Streptoco
14	641.2	56.5	26144	6	AX602184 Sequence
15	641.2	56.5	60266	6	CQ655074 Sequence
16	641.2	56.5	143050	1	SAG766843 Streptoco
17	641.2	56.5	349980	6	CQ655073 Sequence
18	641.2	56.5	349980	6	AX954527 Sequence
19	608.2	53.6	1134	6	AX571491 Sequence

20	608.2	53.6	1137	6	AR480840	AR480840 Sequence
21	608.2	53.6	14418	1	AE007318	AE007318 Streptoco
22	608.2	53.6	21338	6	CQ788929	CQ788929 Sequence
23	608.2	53.6	21338	6	AR218788	AR218788 Sequence
24	608.2	53.6	21338	6	BD003700	BD003700 Polynucle
25	608.2	53.6	232807	2	SPNEU1901	AL449923 Streptoco
26	608.2	53.6	349980	6	AX571766	AX571766 Sequence
27	603.4	53.2	6098	1	SPDNAARG	AF000658 Streptoco
28	603.4	53.2	14222	1	AE008385	AE008385 Streptoco
29	569	50.2	2876	1	AF255728	AF255728 Streptoco
30	408.2	36.0	10186	1	AE006239	AE006239 Lactococc
31	408.2	36.0	10762	1	LLU76424	U76424 Lactococcus
32	339.4	29.9	326434	1	AE016947	AE016947 Enterococ
33	330.6	28.2	8001	6	BD194028	BD194028 Enterococ
34	317	28.0	1155	6	AR346291	AR346291 Sequence
35	252	22.2	343050	1	AL935252	AL935252 Lactobaci
36	247.4	21.8	300787	1	AE016744	AE016744 Staphyloc
37	242	21.3	300007	1	AE017200	AE017200 Lactobaci
38	242	21.3	349980	6	AX926711	AX926711 Sequence
39	236.2	20.8	1131	6	AX621126	AX621126 Sequence
40	236.2	20.8	1134	6	AX554149	AX554149 Sequence
41	236.2	20.8	1134	6	AX741532	AX741532 Sequence
42	236.2	20.8	1134	6	AX741534	AX741534 Sequence
43	236.2	20.8	2347	6	AR354035	AR354035 Sequence
44	236.2	20.8	49336	1	AC025950	AC025950 Staphyloc
45	236.2	20.8	50152	1	AC025948	AC025948 Staphyloc

ALIGNMENTS

RESULT 1	AF280765	Streptococcus pyogenes DNA polymerase III beta subunit (dnaN) gene, partial cds.	1134 bp	DNA	linear	BCT 09-NOV-2001
LOCUS	AF280765	Streptococcus pyogenes DNA polymerase III beta subunit (dnaN) gene, partial cds.	1134 bp	DNA	linear	BCT 09-NOV-2001
ACCESSION	AF280765	Streptococcus pyogenes	1134 bp	DNA	linear	BCT 09-NOV-2001
VERSION	AF280765.2	GI:16885214	1134 bp	DNA	linear	BCT 09-NOV-2001
KEYWORDS	Streptococcus pyogenes					
SOURCE	Streptococcus pyogenes					
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;					
REFERENCE	1 (bases 1 to 1134)	Streptococcus				
AUTHORS	Bruck, I. and O'Donnell, M.					
TITLE	The DNA Replication Machine of a Gram-Positive Organism					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1134)	Streptococcus				
AUTHORS	Bruck, I. and O'Donnell, M.					
TITLE	Direct Submission					
JOURNAL	Submitted (21-JUN-2000)					
REFERENCE	3 (bases 1 to 1134)	DNA Replication, Rockefeller University/HHMI, 1230 York Ave, New York, NY 10021, USA				
AUTHORS	Bruck, I. and O'Donnell, M.					
TITLE	Submitted (09-NOV-2001)					
JOURNAL	University/HHMI, 1230 York Ave, New York, NY 10021, USA					
REMARK	Sequence update by submitter					
COMMENT	On Nov 9, 2001 this sequence version replaced gi:9789548.					
FEATURES	Location/Qualifiers					
source	1..1134					
	/organism="Streptococcus pyogenes"					
	/mol_type="genomic DNA"					
gene	1..1134					
	/db_xref="taxon:1134"					
CDS	1..1134					
	/gene="dnaN"					
	/gene="dnaN"					
	/codon_start=1					
	/transl_table=11					
	/product="DNA polymerase III beta subunit"					
	/protein_id="AAF98349.2"					
	/db_xref="GI:16885215"					

/translation="MIQSFNRLFIHALNTTKRAISKNAIPILSSIKIEVITGT
LTVNGOISIENTIPVSNAGLLITSPAILLEASPFNIISUPDISINVKIEOR
QVVLISGESEITLKGVDQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESR
ILUTGHI VLSNHKDPKAVTDSHRMSQRLITLNDTSADLMVULSKSIRFSAVPTDD
IETVFFSPSGVILFRSEHSIFYLLEGNVPTDRLLMTBFETEVVENTSLRHHME
RAFLSNATQNGTVELEITQNHISAHVNSPEVKYNEDLDIVSQSGDLTISFNPTYL
IESLKAIRSETVKIHFPLSPVRFTLTPGDEBESFQLITPVRN"

ORIGIN

Query Match 100.0%; Score 1134; DB 1; Length 1134;
Best Local Similarity 100.0%; Pred. No. 1.4e-198;
Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60

QY 61 GCTATTAGCACTAAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGTCACTTCT 120
DB 61 GCTATTAGCACTAAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGTCACTTCT 120

QY 121 ACAGGAGTAACCTTTAAGAGTCTTAACGGTCAATATCAATTCGAAACACTATTCCTGTA 180
DB 121 ACAGGAGTAACCTTTAAGAGTCTTAACGGTCAATATCAATTCGAAACACTATTCCTGTA 180

QY 181 AGTAATGAAATGCTGTTTGTCTTAATCTCTCCAGGAGCTATTTTATTAAGAGTCT 240
DB 181 AGTAATGAAATGCTGTTTGTCTTAATCTCTCCAGGAGCTATTTTATTAAGAGTCT 240

QY 241 TTTTATTAATAATTTTCAAGTTTGCAGATATTAGTATTAATTTAAAGAAATGAA 300
DB 241 TTTTATTAATAATTTTCAAGTTTGCAGATATTAGTATTAATTTAAAGAAATGAA 300

QY 301 CAACACCAAGTTCTTTTAAACAGTCTTAATCAGAGTATACCTTTAAAGGAAAGATGT 360
DB 301 CAACACCAAGTTCTTTTAAACAGTCTTAATCAGAGTATACCTTTAAAGGAAAGATGT 360

QY 361 GACAGTATCTCGTCTACAAAGAGTATCAACAGAAATCCTTTGTATTTTAAACAA 420
DB 361 GACAGTATCTCGTCTACAAAGAGTATCAACAGAAATCCTTTGTATTTTAAACAA 420

QY 421 TTATGAGTCTATTTATGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTCTCT 480
DB 421 TTATGAGTCTATTTATGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTCTCT 480

QY 481 ATTTTAAACAGGAGTCTATTTGTTATTAAGTAATCATATAAGATTTTAAAGCAGTAGGACT 540
DB 481 ATTTTAAACAGGAGTCTATTTGTTATTAAGTAATCATATAAGATTTTAAAGCAGTAGGACT 540

QY 541 GACTCTCATCGTATGAGCCACGCTTTAATCACTTTTGACAAATCTTTCAGCAGATTTGATG 600
DB 541 GACTCTCATCGTATGAGCCACGCTTTAATCACTTTTGACAAATCTTTCAGCAGATTTGATG 600

QY 601 GTAGTCTTCCAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATATGAG 660
DB 601 GTAGTCTTCCAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATATGAG 660

QY 661 ACCGTTGAGGATTTTCTCACCAGGCAATCTTTGTTGAGAGTGAACATTTCTTTT 720
DB 661 ACCGTTGAGGATTTTCTCACCAGGCAATCTTTGTTGAGAGTGAACATTTCTTTT 720

QY 721 TATACAGCCTCTTTAGAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAAT 780
DB 721 TATACAGCCTCTTTAGAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAAT 780

QY 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCAGCTATGGAACGCTGCTTCTTG 840
DB 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCAGCTATGGAACGCTGCTTCTTG 840

QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTCAGATTAATCAAAATCATATTTCA 900
DB 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTCAGATTAATCAAAATCATATTTCA 900

QY 901 GCTCATGTTAACTCACTGAGGTTGGTAAGTAAACGAGGATTTAGATATTTAGTACG 960
DB 901 GCTCATGTTAACTCACTGAGGTTGGTAAGTAAACGAGGATTTAGATATTTAGTACG 960

QY 961 TCTGTTAGTATTTAACTATACGCTTCAATCCCACTTACCTTTATTTAGTCTTTTAAAGCT 1020
DB 961 TCTGTTAGTATTTAACTATACGCTTCAATCCCACTTACCTTTATTTAGTCTTTTAAAGCT 1020

QY 1021 ATTTAAAGTGAACAGATAAATAATTCATTTCTTATCACCAGTTTCGACCATTCACCTAACA 1080
DB 1021 ATTTAAAGTGAACAGATAAATAATTCATTTCTTATCACCAGTTTCGACCATTCACCTAACA 1080

QY 1081 CCAGCGATGAGGAAGAAAGTTTATCCAAATTAATTACACAGTACGAAAC 1134
DB 1081 CCAGCGATGAGGAAGAAAGTTTATCCAAATTAATTACACAGTACGAAAC 1134

RESULT 2
LOCUS CQ648158 1134 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 5115 from Patent WO0234771.
ACCESSION CQ648158
VERSION CQ648158.1 GI:41683805
KEYWORDS Streptococcus pyogenes
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE
1 Telford, J., Masignani, V., Margarit y Ros, I., Grandi, G., Fraser, C., and Tettelin, H.
Nucleic acids and proteins from streptococcus groups a & b
Patent: WO 0234771-A 5115 02-MAY-2002;
Chiron S.p.A. (IT); THE INSTITUTE FOR GENOMIC RESEARCH (US)
Location/Qualifiers
1.1134
/organism="Streptococcus pyogenes"
/mol_type="unassigned DNA"
/db_xref="taxon:1314"

ORIGIN

Query Match 99.3%; Score 1126; DB 6; Length 1134;
Best Local Similarity 99.6%; Pred. No. 4.2e-197;
Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60

QY 61 GCTATTAGCACTAAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGTCACTTCT 120
DB 61 GCTATTAGCACTAAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGTCACTTCT 120

QY 121 ACAGAGTAACCTTTTAAACAGGCTTAACGGTCAATATCAATTTGAAACACTATTCCTGTA 180
DB 121 ACAGAGTAACCTTTTAAACAGGCTTAACGGTCAATATCAATTTGAAACACTATTCCTGTA 180

QY 181 AGTAATGAAATGCTGTTTGTCTTAATCTCTCCAGGAGCTATTTTATTAAGAGTCT 240
DB 181 AGTAATGAAATGCTGTTTGTCTTAATCTCTCCAGGAGCTATTTTATTAAGAGTCT 240

QY 241 TTTTATTAATAATTTTCAAGTTTGCAGATTTGCGAGATTTAGTATTAATTTAAAGAAATGAA 300
DB 241 TTTTATTAATAATTTTCAAGTTTGCAGATTTGCGAGATTTAGTATTAATTTAAAGAAATGAA 300

QY 301 CAACACCAAGTTCTTTTAAACAGTCTTAATCAGAGTATACCTTTAAAGGAAAGATGT 360
DB 301 CAACACCAAGTTCTTTTAAACAGTCTTAATCAGAGTATACCTTTAAAGGAAAGATGT 360

QY 361 GACAGTATCTCGTCTACAAAGAGTATCAACAGAAATCCTTTGTATTTTAAACAA 420
DB 361 GACAGTATCTCGTCTACAAAGAGTATCAACAGAAATCCTTTGTATTTTAAACAA 420

QY	421	TTATTGAAGTCTATTATTGCTGAACAGCTTTTGCAGCCAGTATTACAGAAAGTGGTCTCT	480		
Db	421	TTATTGAAGTCTATTATTGCTGAACAGCTTTTGCAGCCAGTATTACAGAAAGTGGTCTCT	480		
QY	481	ATTTTAAACAGGAGTTCATATTGTTAAGTAAATCATAAAGATTTTAAAGCAGTAGCGACT	540		
Db	481	ATTTTAAACAGGAGTTCATATTGTTAAGTAAATCATAAAGATTTTAAAGCAGTAGCGACT	540		
QY	541	GACTCTCATCGTATGAGCCAAACGTTTAACTTTGGACAATACTTTCAGCAGATTTTGATG	600		
Db	541	GACTCTCATCGTATGAGCCAAACGTTTAACTTTGGACAATACTTTCAGCAGATTTTGATG	600		
QY	601	GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTCAGATGATATTGAG	660		
Db	601	GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTCAGATGATATTGAG	660		
QY	661	ACCGTTGAGGTATTTTCTCACCAGGCCAAATCTTGTTCAGAAAGTGAACACATTTCTTTT	720		
Db	661	ACCGTTGAGGTATTTTCTCACCAGGCCAAATCTTGTTCAGAAAGTGAACACATTTCTTTT	720		
QY	721	TATACACGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAATTT	780		
Db	721	TATACACGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAATTT	780		
QY	781	GAGACGGAGGTGTTTTTCAATACCCCAATCCCTTCGCCACGCTATGGAACGTCCTCTTG	840		
Db	781	GAGACGGAGGTGTTTTTCAATACCCCAATCCCTTCGCCACGCTATGGAACGTCCTCTTG	840		
QY	841	ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATTTCA	900		
Db	841	ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATTTCA	900		
QY	901	GCTCATGTTAACTCACTCGAGTTCGTTAAGGTAAACGAGGATTTAGATATTGTTAGTCAG	960		
Db	901	GCTCATGTTAACTCACTCGAGTTCGTTAAGGTAAACGAGGATTTAGATATTGTTAGTCAG	960		
QY	961	TCTGGTAGTGATTTACTATCAGCTTCAATCCAACTTACCTTATGAGTCTTTAAAGCT	1020		
Db	961	TCTGGTAGTGATTTACTATCAGCTTCAATCCAACTTACCTTATGAGTCTTTAAAGCT	1020		
QY	1021	ATTTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCCCTAAC	1080		
Db	1021	ATTTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCCCTAAC	1080		
QY	1081	CCAGCGGATGAGGAAGATTTTATCCAAATTAATACACAGTACGAAACAAAC	1134		
Db	1081	CCAGCGGATGAGGAAGATTTTATCCAAATTAATACACAGTACGAAACAAAC	1134		
RESULT 3	AE006472	12551 bp	DNA	linear	BCT 03-JUN-2004
LOCUS	Streptococcus pyogenes M1 GAS, section 1 of 167 of the complete genome.				
DEFINITION	AE006472	AE004092			
ACCESSION	AE006472.1	GI:13621325			
VERSION					
KEYWORDS	Streptococcus pyogenes M1 GAS				
SOURCE	Streptococcus pyogenes M1 GAS				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.				
REFERENCE	1 (bases 1 to 12551)				
AUTHORS	Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K., Primeaux, C., Sezate, S.S., Surorov, A.N., Kenton, S., Lai, H., Lin, S., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.				
TITLE	Complete genome sequence of an M1 strain of Streptococcus pyogenes				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)				
MEDLINE	21192684				
PUBMED	11296296				
REFERENCE	2 (bases 1 to 12551)				
AUTHORS	Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K., Primeaux, C., Sezate, S.S., Surorov, A.N., Kenton, S., Lai, H., Lin, S.,				

Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.	
Submitted (10-APR-2001) Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 SL Young Blvd, Oklahoma City, OK 73104, USA	
Location/Qualifiers	
1. .12551	
/organism="Streptococcus pyogenes M1 GAS"	
/mol_type="genomic DNA"	
/strain="SP370"	
/serotype="M1"	
/db_xref="taxon:160490"	
232. .1587	
/gene="dnaA"	
/locus_tag="SPY0002"	
232. .1587	
/gene="dnaA"	
/locus_tag="SPY0002"	
/note="Best Blastp hit = gb AAF71535.1 AF255728_1 (AF255728) chromosomal initiator protein DnaA -- (Streptococcus pyogenes]"	
/codon_start=1	
/transl_table=11	
/product="Chromosomal initiator protein"	
/protein_id="AAK31146.1"	
/db_xref="GI:13621326"	
1742. .2878	
/gene="dnan"	
/locus_tag="SPY0003"	
1742. .2878	
/gene="dnan"	
/locus_tag="SPY0003"	
/EC_number="2.7.7.7"	
/note="Best Blastp hit = sp O06672 DP3B_STRPN_DNA POLYMERASE III, BETA CHAIN >gi 2109446 gb AAC45337.1 (AF000658) beta subunit of DNA polymerase III (Streptococcus pneumoniae]"	
/codon_start=1	
/transl_table=11	
/product="beta subunit of DNA polymerase III"	
/protein_id="AAK31147.1"	
/db_xref="GI:13621327"	
/translation="MIOFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVT LTSGNGQISIENTIPVSNENAGLLITSPGAILLEASRPINIISSIPDTSINVKIEIQH OVLITSGKSEITLKGQVDQYPRLOEVSTENPLILKTLKLSIIAETAFAESLQSRP ILTGTHVLSNHKDFKAVATDSHRMSQRLITLDMTSDFDVVIPSKRERSAVFIDD IETVEVFPFSQILFRSEHTSFYRLLEGNYPDTRLMTETEFVETVFNSTLRAME RAFLISSNATQNGTQKLEITQNHISAHVNSPEVKYNDLIDIVSQSGSLTISFNFTYL IESLKATKSETVKIHFILSPVRPFTLTGDEBESFIQLITPVRTN"	
2953. .3150	
/locus_tag="SPY0004"	
2953. .3150	
/locus_tag="SPY0004"	
/note="Best Blastp hit = gb AAC45338.1 (AF000658) ORFX (Streptococcus pneumoniae]"	
/codon_start=1	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="AAK31148.1"	
/db_xref="GI:13621328"	
/translation="MYQIGSFVEMKPKHACVIKETGKXANQWKLVRVGADIKIQCTNC QHVIMMSRYDFERKLKKVLQP"	
3480. .4595	

```

CDS
/locus_tag="SPY0006"
3480..4595
/locus_tag="SPY0006"
/note="Best Blast hit = dbj|BAB07770.1| (AP001520)
GTP-binding protein [Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="AAK33149.1"
/db_xref="GI:13621329"
/translation="MALTAGIVGLPNVGKSTLENAITKAGAEAAAYPPATIDPNVMV
EVPDERLOKLTETLTPKKTPTTFFETDIAGIVKSGREGLGKGNKFLANIREIDAIVH
VRAFDENVMRQGRDAFVDFIADITINLELADLESINKRVARVERKARTOKD
KESAEFNVKQIKPVLKEDGKSARTIEFTEDEAKVVGKFLTLTKFVLVANVEDKV
ANPDGIDYKQIRDFPAFTANAEVAVISARAEEISLDEDEKEEFLAIGLTESGVDK
LTRAAYHLLGLGYFTAGSEKVRMTFKGIIKAPQAQAIHSDFERGFIRAVMTSVD
LMTYGSSEKAVKEAGLREEGKVVQGDIMERFNV"
4665..5234
/gene="pth"
/locus_tag="SPY0007"
4665..5234
/gene="pth"
/locus_tag="SPY0007"
/EC_number="3.1.1.29"
/note="Best Blast hit = dbj|BAB03787.1| (AP001507) stage
V sporulation protein C (peptidyl-tRNA hydrolase)
[Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative peptidyl-tRNA hydrolase"
/protein_id="AAK33150.1"
/db_xref="GI:13621330"
/translation="WMKIVGLGNPGSKYEKTKHNGFMADINIVRNLDVTFDKNF
KAGIGTFINHEKIVFVPTPTMNNISGAVKALLTYNIDITDLIVIVDDLDMEVSK
RLRSGAGCHNGIKSIITAHIGTQFNRIKVGIGRPLKGMTVINHWVGQFETEDNIAI
SITLDRVNVNAVFYLQENDFEXTMQKFN"
5237..8740
/gene="trcf"
/locus_tag="SPY0008"
5237..8740
/gene="trcf"
/locus_tag="SPY0008"
/note="Best Blast hit = gb|AAD03810.1| (AF054624)
transcription-repair coupling factor [Lactobacillus
sakei]"
/codon_start=1
/transl_table=11
/product="putative transcription-repair coupling factor"
/protein_id="AAK33151.1"
/db_xref="GI:13621331"
/translation="WDILFSPQNKVQSWHSGLTTLGRQLVWGLSGSSKTLATASAY
LDOKKIIVVSTQNEVEKLASDLSLDESELVQFFADDDVAAAEFIFASMDKALSRI
ETLQFLRPNKQGVLLIVSLSGLRLLPNPDTFKSQIQLTVGEDYDSDTLTQLMRTIG
YOKVSQVSPGFSRGGDIDYIEITELPYRLEFFGDDIDSIRQFHPETOKSPEOLE
GIFNPASDLIFEVSDFQIEQLEKALQADDDKSYLEDVLAISKNGFKHKDIRF
QSLFYKEWLSLDVTPKGTPIFFDDFOKLVDKNARFLEIANLITLEDLOQKALSNL
YFTDNYRELHKVATPFSNFHKGKLNIPKDMQHLQYAMQEPFQFPLIDIEKRY
OKNOTTVIVQESQVAVRELKSKSQDQVQFRLPLVSAQIVSRESQIVIGASGFYFA
DEKALITEHEIYHKIKRRARRNISNAERLDYNELAVGYYVHNHVGIFGRPLGIE
TIQGIHROIVTYIQNSDRISILPDIQISLSKYVSADGKEPKINKINDGRFQTKQ
KVAQVEDIADLLKLYAERSQKQFSFSPDDQLQRAFDDDFAFVETEDQLRSIKEIK
ADMSQPMQDLVLDVDFGFGTEVAMRAAFKAVNDHKQVAVLPTTVLAQOHHENFKA
RFENYFVVDVLSFRSKQEAETLERVKQIDIIIGTHRLLSKDVVFSDGLIVID
EQRFGVHKETLTKKQDLVLTATPIPTLHMSMLGIRLSVLETFPTNRPVQ
TVLENPLVREALIREMORGGOIFVYNNKVDITIEKKVAELQELVPEASIGFVHGM
SEIQLENTLIDIFNGDDVLVATTIIEGVDISNVNTLFIENADHMGELSTLYQLRGV
GRSNRIATLMDYRDPKLTVEVSEKLEAIKGFTELGSFGFIAMRDLISIRGANGILGA
SQSGFIDSVGTEVMSQLEQAIASKQKTVTRQKNTENIQIDAYLPDDYIADERK
IDIYKRIRESQEDYLNLOELMDRFEGYVDQVAYLLEIALLKHYMDNAFAELVERK
NMQVIRFVETSLTYFLTQDYFEALSNTLHKAKISERHQGKIDIVDFVRHQDYRIIEE
8879..9175
/locus_tag="SPY0009"
8879..9175
/locus_tag="SPY0009"
/note="Best Blast hit = sp|P37557|YABO_BACSU_HYPOTHETICAL
9.7 KDA PROTEIN IN MFD-DIVIC INTERGENIC REGION
>gi|2127042|pir||S66089 conserved hypothetical protein
YABO - Bacillus subtilis >gi|467448|dbj|BAA05294.1|
(D26185) unknown [Bacillus subtilis]
>gi|2632326|emb|CAB11835.1| (Z99104) similar to
hypothetical proteins [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical proteins"
/protein_id="AAK33152.1"
/db_xref="GI:13621332"
/translation="MKRIVGEIMRLDKYLKRSIRLKRSSVAKEVADKGRIVKNGILAK
SSTINLNDHIEISFGNKLTLVRVIEINDSTKKEDALKMYEIISETRITLNEEA"
9162..9533
/gene="divic"
/locus_tag="SPY0010"
9162..9533
/gene="divic"
/locus_tag="SPY0010"
/note="Best Blast hit = gb|AAC98903.1| (AF023181) Divic
homolog [Listeria monocytogenes]"
/codon_start=1
/transl_table=11
/product="putative cell division protein (Divic)"
/protein_id="AAK33153.1"
/db_xref="GI:13621333"
/translation="MKKPSIVQLNNHYIKENLKKPEEESOKRNFMGWILVSMVF
LFILPTYNLVKSYVDPEKQVVKLKKYNEELSESTKKEQLAERLKDNDNFVKYAR
AKYLSREGEIMYPIPLGLPK"
Query Match 99.3%; Score 1126; DB 1; Length 12551;
Best Local Similarity 99.6%; Pred. No. 2.6e-197;
Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGATTCAAATTTTCAATTAATCGCACATTATTTATTCATCGCTTTTAAATACAACTAAACCGT 60
DB 1742 ATGATTCAAATTTTCAATTAATCGCACATTATTTATTCATCGCTTTTAAATACAACTAAACCGT 1801
QY 61 GCTATTAGCACTAAAATGCAATTCCTATTCTTTTCATCAATAAAATTTGAAGTCACCTCT 120
DB 1802 GCTATTAGCACTAAAATGCAATTCCTATTCTTTTCATCAATAAAATTTGAAGTCACCTCT 1861
QY 121 ACAGGAGTAACTTTAACAGGCTCTAACGGTCAATATCAATTCGAAACACTATTCCTGTGTA 180
DB 1862 ACAGGAGTAACTTTAACAGGCTCTAACGGTCAATATCAATTCGAAACACTATTCCTGTGTA 1921
QY 181 AGTAATGAAAATGCTGGTTTGTGTAATTAATCCTCTCCAGGAGCTATTTATTAGAGAGCTAGT 240
DB 1922 AGTAATGAAAATGCTGGTTTGTGTAATTAATCCTCTCCAGGAGCTATTTATTAGAGAGCTAGT 1981
QY 241 TTTTATTATTAATTAATTTCAAGTTTGCAGGATATTAGTATAAATGTTTAAAGAAATTTGAA 300
DB 1982 TTTTATTATTAATTAATTTCAAGTTTGCAGGATATTAGTATAAATGTTTAAAGAAATTTGAA 2041
QY 301 CAACACCAAGTTGCTTTTAAACAGTGTGTAATCAGAGATTACCTTTAAAGAGAAAGATGTT 360
DB 2042 CAACACCAAGTTGCTTTTAAACAGTGTGTAATCAGAGATTACCTTTAAAGAGAAAGATGTT 2101
QY 361 GACAGTATCCTCGTCTCAAGAAGTATCAACAGAAAATCCTTTGATTGTTTTAAAAACAAAA 420
DB 2102 GACAGTATCCTCGTCTCAAGAAGTATCAACAGAAAATCCTTTGATTGTTTTAAAAACAAAA 2161
QY 421 TTATTCAAGTCTATTATTCGTAACAGCTTTTGACGCGAGTTTACAGAAAGTCCTCT 480
DB 2162 TTATTCAAGTCTATTATTCGTAACAGCTTTTGACGCGAGTTTACAGAAAGTCCTCT 2221
QY 481 ATTTTAAAGGAGTTCATATTTGTTAAGTAATCAATAAGATTTTAAAGCAGTAGGACT 540
DB 2222 ATTTTAAAGGAGTTCATATTTGTTAAGTAATCAATAAGATTTTAAAGCAGTAGGACT 2281

```


Qy	541	GACTCTCATCGTATGAGCCAGGTTTAATCACTTTGGACAACTCTTGACAGATTTGATG	600
Db	2282	GACTCTCATCGTATGAGCCAGGTTTAATCACTTTGGACAACTCTTGACAGATTTGATG	2341
Qy	601	GTAGTTCTTCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTCAG	660
Db	2342	GTAGTTCTTCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTCAG	2401
Qy	661	ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTTGTCAGAGTGAACACATTTCTTTT	720
Db	2402	ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTTGTCAGAGTGAACACATTTCTTTT	2461
Qy	721	TATACACGCTCTTTAGAGGAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT	780
Db	2462	TATACACGCTCTTTAGAGGAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT	2521
Qy	781	GAGCGGAGGTGTTTTCACCAACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTG	840
Db	2522	GAGCGGAGGTGTTTTCACCAACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTG	2581
Qy	841	ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAAATCATATTTCA	900
Db	2582	ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAAATCATATTTCA	2641
Qy	901	GCTCATGTTAACTCACTGAGTTCGTAAGGTAAACGAGGATTTAGATATTTGTTAGTCAG	960
Db	2642	GCTCATGTTAACTCACTGAGTTCGTAAGGTAAACGAGGATTTAGATATTTGTTAGTCAG	2701
Qy	961	TCTGTAGTGATTAACTATCAGCTTCAATCAACTTACCTTATTTAGCTCTTTAAAGCT	1020
Db	2702	TCTGTAGTGATTAACTATCAGCTTCAATCAACTTACCTTATTTAGCTCTTTAAAGCT	2761
Qy	1021	ATTTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCCTAACA	1080
Db	2762	ATTTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCCTAACA	2821
Qy	1081	CCAGCGATGAGGAAGTATTTTATCAATTAATTAATCACCAGTACGACCAAC	1134
Db	2822	CCAGCGATGAGGAAGTATTTTATCAATTAATTAATCACCAGTACGACCAAC	2875

RESULT 4

CP000003.00

WPCOMMENT

Sequence split into 19 fragments LOCUS CP000003 Accession CP000003

Fragment Name	Begin	End
CP000003_00	1	110000
CP000003_01	100001	210000
CP000003_02	200001	310000
CP000003_03	300001	410000
CP000003_04	400001	510000
CP000003_05	500001	610000
CP000003_06	600001	710000
CP000003_07	700001	810000
CP000003_08	800001	910000
CP000003_09	900001	1010000
CP000003_10	1000001	1110000
CP000003_11	1100001	1210000
CP000003_12	1200001	1310000
CP000003_13	1300001	1410000
CP000003_14	1400001	1510000
CP000003_15	1500001	1610000
CP000003_16	1600001	1710000
CP000003_17	1700001	1810000
CP000003_18	1800001	1899877

LOCUS CP000003 1899877 bp DNA circular BCT 03-AUG-2004

DEFINITION Streptococcus pyogenes MGAS10394, complete genome.

ACCESSION CP000003

VERSION CP000003.1 GI:50902420

KEYWORDS

SOURCE Streptococcus pyogenes MGAS10394

ORGANISM Streptococcus pyogenes MGAS10394

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Streptococcus.
1 (bases 1 to 1899877)
Banks,D.J., Porcella,S.F., Barbican,K.D., Beres,S.B., Philips,L.E.,
Voyich,J.M., DeLeo,F.R., Martin,J.M., Somerville,G.A. and
Musser,J.M.
Progress toward Characterization of the Group A Streptococcus
Metagenome: Complete Genome Sequence of a Macrolide-Resistant
Serotype M6 Strain
J. Infect. Dis. 190 (4), 727-738 (2004)
15272401
2 (bases 1 to 1899877)
Banks,D.J., Porcella,S.F., Barbican,K.D., Beres,S.B., Philips,L.E.,
Voyich,J.M., DeLeo,F.R., Martin,J.M., Somerville,G.A. and
Musser,J.M.
Direct Submission
Submitted (20-JUL-2004) Laboratory of Human Bacterial Pathogenesis,
Rocky Mountain Laboratories, National Institute of Allergy and
Infectious Disease, National Institutes of Health, 903 S. 4th
Street, Hamilton, MT 59840, USA
Location/Qualifiers
1. 1899877
/organism="Streptococcus pyogenes MGAS10394"
/mol_type="genomic DNA"
/strain="MGAS10394"
/serotype="M6"
/db_xref="taxon:286636"
202. 1557
/locus_tag="M6_Spy0001"
202. 1557
/locus_tag="M6_Spy0001"
/note="Dnaa"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein"
/protein_id="AAT86136.1"
/db_xref="GI:50902421"
/translation="MTENEQIFWNRVLELAQSOLKQATYEFFVHDARLLKVDKHIATI
YLDQMKELFWKNUKDVLTAGFEVYNAQISVDVFEDLMIEQNTKINQPKQOAL
NSLPTVTDLSNYSFENFIQGDENRWAASIAVANTPGTYYNPLFIWGGFGLKTH
LLNAIGNSVLLNPNARIKYITAEFNIFEVTHIRLDTMDLDELKERLNDLLLIIDIO
SLAKKTLSGTOEEFFNTFNALNNKQIVLTSDRTPDHLNDELKRLVTRFKWGLTVNI
TPDPFETRVAILTNKIOEYNFIFPDQTIYLAGOFDSNVRLDGLAKDISLVANFKOI
DTIVFAAEAIRARKQDGPMTVIPBEEIOAGKFGYGVTVKEIKAKRTQNVILAR
QVAMFLAREMTDNLSPKIGFEFGGRDHSVLHAYNKIKNMISQDESRIETIKNKI
K"
1712. 2848
/locus_tag="M6_Spy0002"
1712. 2848
/locus_tag="M6_Spy0002"
/EC_number="2.7.7.7"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta chain"
/protein_id="AAT86137.1"
/db_xref="GI:50902422"
/translation="MIQFSINRTLFIHALNATKRAISTKNAIPILSSIKIEVTPGTVT
LTGNSQISIENTIPVSNENAGLLITSPGAILLEASFNIISSLPDISINVKETEIQH
QVLTSGKSEITLTKGVDDQYPRLQEVSTENPLIKTKLKSIIAETAFASLQESRP
ILTGHVLSNHKDFKPAHSDHMSORLITLDTNSADPDVVPISLRFSAVFTDD
IETVEVFPSPQILFRSEHISFYTRLLGNYPTDRLMLTFETEVFTQSLRHAME
RAFLSNATQNGTVKLEITQNHISAHVNSPEVGKVNEDLDIVSQSGSLTISFNPYTL
IESUKAIRKETVKIHFUSPVPFPFTLTGDBEESFIQLITPVRTN"
2923. 3120
/locus_tag="M6_Spy0003"
2923. 3120
/locus_tag="M6_Spy0003"
/codon_start=1
/transl_table=11
/product="Hypothetical cytosolic protein"
/protein_id="AAT86138.1"
/db_xref="GI:50902423"
/translation="MYQIGSFVEMKKPHACVIKETGKKANQMKVLRVGADIKIQCTNC
QHVIMSRYDFDRKKVLQP"

```
gene
3450..4565
/locus_tag="M6_Spy0004"
CDS
3450..4565
/locus_tag="M6_Spy0004"
/notes="probable translation factor"
/codon_start=1
/transl_table=11
/product="GTP-binding protein"
/protein_id="AA086139.1"
/db_xref="GI:50902424"
/translation="MALTAGIVGLPNVKGSTLFLNATKAGAAANYPFATIDPNVGMV
EVPDRLOKLTETLTPKTVPTPEETDIDTIVKASRGEGLNKFLANIREVDALVH
VWPAEDENVMREOGREDAPVDIADIDITINLELILADLESINKRYARVERKWARQKD
KVSABFNVLQKIPVLEDEKGARTIEFTEBAKVKVGLFLLTTPFLVLYVAVNVDKDV
ANPDGIDYVQKIRDAENAEVVIISARABEEISLDEDEKEEFLEAIGFTESGVDK
LTRAAYHLGLGTFTAGKEVRAMTFKRGIKAPQAAGIHSDFERGFIRAVTMSYDD
LMIYSEKAVKAGRLREGEKGVVQDGDIMEFRNV"
4590..5204
/locus_tag="M6_Spy0005"
gene
4590..5204
/locus_tag="M6_Spy0005"
/EC_number="3.1.1.29"
/codon_start=1
/transl_table=11
/product="Peptidyl-tRNA hydrolase"
/protein_id="AA086140.1"
/db_xref="GI:50902425"
/translation="MEEHSSPFGILERNVMKIVGLGNPGSKYKETHKHNIGFMAIND
IVNLDVPTDDKFNKAQIGFTFINHEKVYFVKPTTFNNNSGIAVKALLTYVINDIDT
LVIYDLDMEYSKLRLSKSGAGHGKISIIAHIGTGFENRINKVIGRPLKGMTVI
SHVMQGFNTEDNIALSLTLDVVNAVKPYLQENDPEKTMQKFG"
5207..8710
/locus_tag="M6_Spy0006"
gene
5207..8710
/locus_tag="M6_Spy0006"
/codon_start=1
/transl_table=11
/product="transcription-repair coupling factor"
/protein_id="AA086141.1"
/db_xref="GI:50902426"
/translation="MDILELFSQNKVQSHSGTLTLGRQLVMGLSGSSKALATASAV
LDDQKIVVSTQNEVKLASLSDLEELVQFPADDDVAEAEEFTPASMDKALSRI
ETLQPLRNPQSGVLIVSLSLTLNPNPDVFTKSOIOLTWGEDYDSDTLQKMLTIG
YKQSVQVISPGEFRGDIIDIVETQELPRLRLEFFGDDIDSIOQFTPELTKSEFQLE
GIFINPASDLIFASDPFRGIEQLEKQTAQDDKKSYLEDLVAVSKNGFKHKDIRF
QSLFYEKWSLDDYIPKGTPIFFDDFKLVKNARFDLEIANLLTELQOQKSNLN
QKQNTVIVQVESQVAYERLEKSFQDYQFRLPLVSAQIVRSQIVIGATSSGFYFA
DEKALITHEIYHKIKRRARRNSINAEKLDYNEIYAVGQVYVHNVHVGIGRFLGE
TIOQIHRDVVTIYQONSDRISLPIDQIGSLSKYVSADGKEPKINKLNGRFOKTKQ
KVAREVIDADLLKLYAERSQQKGFSPDDDLQAPDDDFAFVETEDQLRSIKEIK
ADMSQMQRLLVDVGFQKTEVAMRAAFKAVNDHQVAVLVTPTVLAQQHYENFKA
RFENYPVEVDVLSRFSKKEQAETLERVQKQDIDIIIGTHRLSKDVSFSLGILVID
BEQYFGVKHETLKLTKVDLVLTATPIPTLHSMGLGIRLSVETPTPNRYPVQ
TYVLNPLGVREAIIEEMDRGGQIFVYVANKVDITIEKKVAELQELVPEASIGFVHGOM
SEIOLENTLIDFINGVDVLVATTIETGVDISNVTIETENADHMGSLTYQLRGVR
GRSRIAYAVLMVYPKDILTEVSEKRLKLEAKGFTELSGFKIARMDLSIRGAGNIGA
SQSGFIDSVGEMYSQLLLEQAIASQKQTVVRKQNTIEINLQIDAYLPDDYIADERQK
INDYKIREISRESDYNLQDELIDFEGEPDQVAYLLEIALLKHYMDNAPAEIIVERK
NNQVIREFVSTSLTYFLTDQYFELSCKTHLKAKISEHQGKIDIVFDRVHQDVRILBE
LMLFGERLSEIKIRKNSVFK"
8848..9144
/locus_tag="M6_Spy0007"
gene
8848..9144
/locus_tag="M6_Spy0007"
/codon_start=1
/transl_table=11
/product="Heat shock protein 15"
/protein_id="AA086142.1"
/db_xref="GI:50902427"
/translation="MKRIVGEIMRLDKYLKVSRLKRRSVAKEVADKGRIKVNGIILAK
SSINVKLNDSHIEISFGNKLITVRVIEIKDKSKEDALKWYIEIISRTITLNEEA"
9131..9502
gene
/locus_tag="M6_Spy0008"
9131..9502
/locus_tag="M6_Spy0008"
/notes="cell division protein"
/codon_start=1
/transl_table=11
/product="DivIC"
/protein_id="AA086143.1"
/db_xref="GI:50902428"
/translation="MKKPSIVOLNHHYIKKENLKKKFFEEBSOKRNRFRMGWILVSMWF
LFIPLTYNLVKSVDYDFEKQNOQVVKLKKKEYNELSKTKKERQLAERLKKDDNFVKYAR
AKYLSRREGIMYIPGLLPK"
9499..9624
/locus_tag="M6_Spy0009"
gene
9499..9624
/locus_tag="M6_Spy0009"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AA086144.1"
/db_xref="GI:50902429"
/translation="MMMDNIIKKIEAFIAPSFDDKKLAEQENQKVEESQYIQKN"
9637..10923
/locus_tag="M6_Spy0010"
Query Match
Best Local Similarity 98.6%; Score 1118; DB 1; Length 110000;
Matches 1124; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 ATGATTCAATTTTCAATTAATCGCACATTAATTCATGCTTTAATGCACTAAACGT 60
Db 1712 ATGATTCAATTTTCAATTAATCGCACATTAATTCATGCTTTAATGCACTAAACGT 1771
Qy 61 GCTATTAGCACTAAATGCAATTCCTATTCTTTCATCAATAAATAAGTCACTTCT 120
Db 1772 GCTATTAGCACTAAATGCAATTCCTATTCTTTCATCAATAAATAAGTCACTTCT 1831
Qy 121 ACAGGAGTAACTTTAAACAGGCTTAACGGTCAATATCAATTTGAAACACATTTTCCCTGTA 180
Db 1832 ACAGGAGTAACTTTAAACAGGCTTAACGGTCAATATCAATTTGAAACACATTTTCCCTGTA 1891
Qy 181 AGTAATGAAATGCTGTTTGGTAAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
Db 1892 AGTAATGAAATGCTGTTTGGTAAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 1951
Qy 241 TTTTATTAAATATTAATTTCAAGTTTGCAGATATTAGTATAATGTTAAAGAAATTTGAA 300
Db 1952 TTTTATTAAATATTAATTTCAAGTTTGCAGATATTAGTATAATGTTAAAGAAATTTGAA 2011
Qy 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAGATGTT 360
Db 2012 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAGATGTT 2071
Qy 361 GACCACTATCTCGTCTACAGAAAGTATCAACAGAAATCTCTTGTGATTTTAAAAACAAA 420
Db 2072 GACCACTATCTCGTCTACAGAAAGTATCAACAGAAATCTCTTGTGATTTTAAAAACAAA 2131
Qy 421 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCT 480
Db 2132 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCT 2191
Qy 481 ATTTTAAACAGGAGTTTCAATTTGATTAAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
Db 2192 ATTTTAAACAGGAGTTTCAATTTGATTAAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 2251
Qy 541 GACTCTCATCGTATGACCAAGCTTTTAATCACTTTGACAACTATTTTACAGAGATTTTATG 600
Db 2252 GACTCTCATCGTATGACCAAGCTTTTAATCACTTTGACAACTATTTTACAGAGATTTTATG 2311
Qy 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTGAG 660
Db 2312 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTGAG 2371
```

QY 661 ACCGTTGAGTATTTTCTCACAAGCCAAATCTTGTTCAGAGTGACACATTTCTTTT 720
Db 2372 ACCGTTGAGTATTTTCTCACAAGCCAAATCTTGTTCAGAGTGACACATTTCTTTT 2431
QY 721 TATACACGCTCTTAGAGGAATATATCCGATACAGACCGTTTATTAATGACAGATTT 780
Db 2432 TATACACGCTCTTAGAGGAATATATCCGATACAGACCGTTTATTAATGACAGATTT 2491
QY 781 GAGACGGAGTTGTTTCAATACCAATCCCTTCGCCAGCTTATGGAACGTCCTTCTTG 840
Db 2492 GAGACGGAGTTGTTTCAATACCAATCCCTTCGCCAGCTTATGGAACGTCCTTCTTG 2551
QY 841 ATTTCTAATGCTACTCAAAATCGTACTGTTAGCTTGAGATTACTCAAAATCATATTTCA 900
Db 2552 ATTTCTAATGCTACTCAAAATCGTACTGTTAGCTTGAGATTACTCAAAATCATATTTCA 2611
QY 901 GCTCATGTTAACTCACCTGAGCTTGGTAAAGTAAACGAGGATTTAGATATTCTTAGTCAG 960
Db 2612 GCTCATGTTAACTCACCTGAGCTTGGTAAAGTAAACGAGGATTTAGATATTCTTAGTCAG 2671
QY 961 TCTGTTAGTGAATTTAACTATACGTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 1020
Db 2672 TCTGTTAGTGAATTTAACTATACGTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 2731
QY 1021 ATTAAGAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCATTCACCCCTAACA 1080
Db 2732 ATTAAGAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCATTCACCCCTAACA 2791
QY 1081 CCAGCGGATGAGGAAGAAAGTTTATCCAAATTAATACACAGTACGACCAAC 1134
Db 2792 CCAGCGGATGAGGAAGAAAGTTTATCCAAATTAATACACAGTACGACCAAC 2845

RESULT 5
LOCUS AE0009953 14943 bp DNA linear BCT 03-APR-2002
DEFINITION Streptococcus pyogenes strain MGAS8232, section 1 of 173 of the complete genome.
ACCESSION AE0009953 AE009949
VERSION AE009953.1 GI:19747290
KEYWORDS
SOURCE
ORGANISM Streptococcus pyogenes MGAS8232
Streptococcus pyogenes MGAS8232
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 14943)
Smoot, J.C., Barbican, K.D., Van Gompel, J.J., Smoot, L.M., Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M., Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M., Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G. and Musser, J.M.
Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks
Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)
21927593
11917108
MEDLINE
PUBMED
REFERENCE
AUTHORS Smoot, J.C., Barbican, K.D., Van Gompel, J.J., Smoot, L.M., Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M., Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M., Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G. and Musser, J.M.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Laboratory of Human Bacterial Pathogenesis/Rocky Mountain Laboratories/NIH, 903 S. 4th St., Hamilton, MT 59840, USA
FEATURES
source
1. .14943
/organism="Streptococcus pyogenes MGAS8232"
/mol_type="genomic DNA"
/strain="MGAS8232"
/db_xref="taxon:186103"
202. .1557
/gene="dnaA"

/note="synonym: spyM18_0001"
202. .1557
/gene="dnaA"
/note="best blastp match sp|Q9L572|DNAA_STRPY CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein"
/protein_id="AAL96837.1"
/db_xref="GI:19747291"
/translation="MTNEQIQFVNRVLEAQSOLKQATYEFFVHDARLLKVDKHIATII YLDOMKELFEKNLKDVLITAGFEVYNAQISVDYVFREDLMEQNTQINOKPKQOAL NSLPTVSDLSKYSFENFIQDENDNRWAASIAVANTPGTYTNPLFWGSPGLKTH LNAIGNSVLLENENARIKIYTAENFINFVHIRLDMDELKEFRNDDLLDIDDIQ SLAKTLSTGTOEEFFNTFNALHNNKQIVLTSDRTPDHLNDELRLVFRFKWGLTVNI TTPDFEFTVAITNKKIIOYFNIFPDQTIETLAGOFDSNVRLGALKQISLVANFKQI DTITVDIAAEAIRAKODGPKMTVPIEBEIOAQVGKFGYVTVKEIKATKRTQNIVLAR QVAMFLAREMTDNLSPKIGKEFGGRDHSVTLHAYNKIKNMISQDESLEIETIKNKI K"
1709. .2845
/gene="dnaN"
/note="synonym: spyM18_0002"
1709. .2845
/gene="dnaN"
/note="best blastp match gb|AAK33147.1| (AE006472) beta subunit of DNA polymerase III (Streptococcus pyogenes M1 GAS)"
/codon_start=1
/transl_table=11
/product="beta subunit of DNA polymerase III"
/protein_id="AAL96838.1"
/db_xref="GI:19747292"
/translation="MIQFSINRLLFIHALNATKRAISTKNAIPILSSIKIKVPTGVT LTSGNQTSIENTIPVSNENAGLLITSPGAILLEASFNIISSLPDISINVKEIEQH QWLTSGKSEITLKGKVDQYPRQVSTENPLILKTLKLSIIAETAPASLOQSRP ILTGHVILSHNKKDKAVATDSHMSORLITLDMTADFDVVPKSRRESAVTDD IETVEFPSPQILFRSEHISFYRLLEGVYPTDRLMLTEFEVWVNTQSLRAME RAFLISNATQNGTVKTLQNHISAHVNSPEVGNVDLIDIVSQGSOLITISFNPTYL TESLKAITSKTVKIFLSPVRPFTLTGDEBESFIQLITPVRTN"
3447. .4562
/gene="spyM18_0004"
3447. .4562
/gene="spyM18_0004"
/note="best blastp match gb|AAK33149.1| (AE006472) putative GTP-binding protein (Streptococcus pyogenes M1 GAS)"
/codon_start=1
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="AAL96839.1"
/db_xref="GI:19747293"
/translation="MALTAGIVGLPNVGKSTLFNAITKAGAEAAANYPPATIDPNVGMV EYPERLQKLTETLPKKTVPFTTFTDIAGVKGASRGELGNKFLANIREVDAIVH VVRAPDENVMREOGRDADFVDIADITINLEILADLESINKRYARVEKMWARTQK KESVAENVLQKIRPLEVDGKSARTIETEEEEKVKVKGFLITTKPVLYVYANVEDKY ANPDGIDYVKQIRDFAATENAENVVISARABEEISELDEDEKFELEALGFTESGVDK LTRAYHLLGLGTYFTAGKEVRWMTFKRGIKAPQAAGIIHSDFERGFIRAVTMSYDD LMIYGSKAVKAGRLREGKEYVYVQGDIMEFRFNV"
4632. .5201
/gene="pth"
/note="synonym: spyM18_0005"
4632. .5201
/gene="pth"
/note="best blastp match gb|AAK33150.1| (AE006472) putative peptidyl-tRNA hydrolase (Streptococcus pyogenes M1 GAS)"
/codon_start=1
/transl_table=11
/product="putative peptidyl-tRNA hydrolase"
/protein_id="AAL96840.1"
/db_xref="GI:19747294"
/translation="MVKMIYGLGNPGSKYEKTKUNIGFMAIDNIVKNDLVDTDDKNF KAQIGSTFINHEKVFVKPTFMNNSGIAVKALLTYNIDITDLIIVYDDLDMEVSKL

```

RLRSKSGAGHNGIKSIIAHIGTQEFNRKIVGIRPLKGMVTI5HVWQFNTEDNIAI
SLTLDRVNVNAVKFLQENDFEKTMQKENG"
/ gene 5204. .8707
/ note="trcf"
/ gene="trcf"
/ synonym: spyM18_0007"
/ CDS 5204. .8707
/ gene="trcf"
/ note="best blastp match gb|AAK33151.1| (AE006472)
putative transcription-repair coupling factor
[Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transl_table=11
/ product="putative transcription-repair coupling factor"
/ protein_id="AAL96841.1"
/ db_xref="GI:19747295"
/ translation="MDILELFSQNKVQSWHSLGTLGRQLVWGLSGSKALAIASAY
LDOKKIVVTVSTQNEVEKLASDLSSLDLVEFPQADVAARAEFASMDKALSRI
ETLQFLNPKSQGLVLSLSLRLPNDPFTKSIQUTVGEDYDSDTLTKQLMTIG
YKVSQVISPGEFSRGGDILDIYEITQELPYRLEFGDDISIRQPHPTQKSFQLE
GIFINPADDIFASDFQRIQLEKALQTAQDNKSYLDELAVLAVKNGFKHKIRKF
QSIFYKEWLLDYIPKGTPIFFDDFKLVKNARPDLEIANLLTDLQOQKALNKL
YFADNRELSHYKPAIFSNFKGLGNIKFQDWHOLTOYMOEFENQPELLDEIKRY
QKQNTTVIVQESQYAYERLKSFOYQFRLPLVLSANOIVRSQIIVGAISGFFA
DDKALITHEIYHKIKRARSNSIAERLKNELAVGDIYVHNHVGIGRFIGIE
TIOGHRDYVTIOYNSDRISLPDIQISLSKVSADGKPKINKLNDGRFQTKQ
KVARVEDTADLLKLYAERSQKGFSPDDDLQAFDDDFAFVETEDQLRSIKEIK
ADNESQPMRDLVGVGFGKTEVAMRAAFKAVNOHKQVVLVPTTVAQOQHYENFKA
RPNYVEVDVLSRFSKQEAETLVERVKQDIDIIIGTHRLLSKDVVFSFDPLVID
EORQVFKHETLKEIKTKVDLTATPIPTLHMSMLGIRLSVITPTNRPVQ
TVLENNPGLREAIIREMDRGQVYVYVNVKDTIDKKVAELQELVPEASIGFVHQGM
SEIQLNTLIDFNGDYDVLVATTIETGVDISNVNTPIENADHMGSLTYQLGRV
GRSNRTAYAVLPDKVLTVESEKLEIKGFTELSGGFKIAMDRLSIRGAGNIIA
SOSGFTDVGPMYSQQLLEQASKOGKTTVRKQNTINQIDAYLDDYISBERQK
IDYKRIKQREDYLNQLEMDRFGQYDPQVAYLLEIALIKHMDNTFAELVERK
NQNVIRFVTSLSYPLTDQYFEALSKHLKAKISEHQKIDIVFDVRHQDYRLIE
8869. .9141
/ gene="spyM18_0008"
8869. .9141
/ gene="spyM18_0008"
/ note="best blastp match gb|AAK33152.1| (AE006472)
conserved hypothetical protein [Streptococcus pyogenes M1
GAS]"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="AAL96842.1"
/ db_xref="GI:19747296"
/ translation="MRLDKYLVKSRLIKRRSVAKVADKGRIVKNGILAKSSTNVKLN
DHIEIFGNKLLTVRVIBIKDSTKEDALKMVEIISETRITLNEA"
9128. .9499
/ gene="divic"
/ note="synonym: spyM18_0009"
9128. .9499
/ gene="divic"
/ note="best blastp match gb|AAK33153.1| (AE006472)
putative cell division protein (DivIC) [Streptococcus
pyogenes M1 GAS]"
/ codon_start=1
/ transl_table=11
/ product="putative cell division protein"
/ protein_id="AAL96843.1"
/ db_xref="GI:19747297"
/ translation="MKKPSIVQLNNHYIKKENLKKPFEESOKRNFEMGILVSNMF
LFILPTNVLKSVYDVPKQVVKQVVKKEYNELSKTKKEQLAERLKDNDVFKVKAR
AKVYLSREGMIYPIPLLPK"
9634. .10920
/ gene="spyM18_0011"
9634. .10920
/ gene="spyM18_0011"
/ note="best blastp match gb|AAK33154.1| (AE006472)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/ codon_start=1

```

```

/ transl_table=11
/ product="hypothetical protein"
/ protein_id="AAL96844.1"
/ db_xref="GI:19747298"
/ translation="MRKJLAAMLMTFFLPLPVIKSTEKKLIFSNAVYQLKQDVVQST
QFYNQPSNPNIYQETCAKSDSLTLPAGRLGVNQPLLKSLVNLKSLPVELADGT
YEANRQLIYDDIVLNQVDIDSYFWTKKRLYSAPYVLGTQITPSPSPAKVHATQ
MAQTNHGYIYIDDKWASQBDLVQFONRMLKVOEMLLOKNNPNYSIFVKOLNTQIS
AGINADKQYAAASIKLAPYIVOKLQKKLAENKTLTYTKDVNHYGYDYPDLOGSK
ISKIADNKDYREDLLKAVAOQSDNATNLYLYLCHOYDRAFRSEIKALSGIDWME
ORLLTSRSANMEALHQQKQIISYLSNTEFDOORITKNITVPVAKHIGDAYDKHD
VAIYGNTPFSLISFTNKSDIITADDDVYILK"
10917. .12203
/ gene="mesj"
/ note="synonym: spyM18_0012"
10917. .12203
/ CDS
Query Match 98.4%; Score 1116.4; DB 1; Length 14943;
Best Local Similarity 99.0%; Pred. No. 1.4e-195;
Matches 1123; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGATTCAAATTTTCAATTAATCGCACATTAATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 1709 ATGATTCAAATTTTCAATTAATCGCACATTAATTTATTCATGCTTTTAAATACAACTAAACGT 1768
QY 61 GCTATTAGCACTAAAATGCCATTCTTCTATTCTTTTCATCAATAAAAATTGAAGTCACCTCT 120
Db 1769 GCTATTAGCACTAAAATGCTATTCTTCTATTCTTTTCATCAATAAAAATTGAAGTCACCTCT 1828
QY 121 ACAGAGTAACCTTTAAACAGGGTCTAACGGGTCAATATCAATTCATCAATAAAAATTGAAGTCACCTCT 180
Db 1829 ACAGAGTAACCTTTAAACAGGGTCTAACGGGTCAATATCAATTCATCAATAAAAATTGAAGTCACCTCT 1888
QY 181 AGTAATGAAATGCTGGTTGCTAAATTAATCTCTCCAGGAGCTATTTTATTAGAGCTAGT 240
Db 1889 AGTAATGAAATGCTGGTTGCTAAATTAATCTCTCCAGGAGCTATTTTATTAGAGCTAGT 1948
QY 241 TTTTATTAAATATTTTCAAGTTTCCAGATATTAGTATAAATGTTTAAAGAAATGAA 300
Db 1949 TTTTATTAAATATTTTCAAGTTTCCAGATATTAGTATAAATGTTTAAAGAAATGAA 2008
QY 301 CAACACCAAGTTGTTTAAACAGGTGTAATTAATCACTCTCCAGGAGCTATTTTAAAGAAAGATGTT 360
Db 2009 CAACACCAAGTTGTTTAAACAGGTGTAATTAATCACTCTCCAGGAGCTATTTTAAAGAAAGATGTT 2068
QY 361 GACCATGATCCTGCTCTACAGAGTATCAACAGAAATCTTGTGATTTTAAAGAAAGATGTT 420
Db 2069 GACCATGATCCTGCTCTACAGAGTATCAACAGAAATCTTGTGATTTTAAAGAAAGATGTT 2128
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGGCAGCCAGTTTACAGAAAGTCTGCT 480
Db 2129 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGGCAGCCAGTTTACAGAAAGTCTGCT 2188
QY 481 ATTTTAAAGGAGTTCATATTTGATTAAGTAAATATAAGATTTTAAAGCAGTACGACT 540
Db 2189 ATTTTAAAGGAGTTCATATTTGATTAAGTAAATATAAGATTTTAAAGCAGTACGACT 2248
QY 541 GACTCTCATGCTATGAGCCCAAGCTTTTAAAGTAAATTTTTCAGAGTATTTTTCAGAGATTTTGA 600
Db 2249 GACTCTCATGCTATGAGCCCAAGCTTTTAAAGTAAATTTTTCAGAGTATTTTTCAGAGATTTTGA 2308
QY 601 GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGAGTATTTTTCAGAGTATTTTTCAG 660
Db 2309 GTGGTTTATTCAGTAAATCTTTGAGAGAAATTTTTCAGAGTATTTTTCAGAGTATTTTTCAG 2368
QY 661 ACCGTGTAGGTATTTTCTCCAGGAGCAATCTTGTTCAGAGTAAACACATTTCTTTT 720
Db 2369 ACCGTGTAGGTATTTTCTCCAGGAGCAATCTTGTTCAGAGTAAACACATTTCTTTT 2428
QY 721 TATACACGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTCTTATTAATCAGACAAATTT 780
Db 2429 TATACACGCTCTTAGAAGGAAATTTATCCCGATACAGACCGTCTTATTAATCAGACAAATTT 2488

```

QY 781 GAGCGGAGGTGTTTCAATACCAATCCCTTCGCCAGCTATGGAAGTGCCTCTCTTG 840
Db 2489 GAGCGGAGGTGTTTCAATACCAATCCCTTCGCCAGCTATGGAAGTGCCTCTCTTG 2548
QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATTCA 900
Db 2549 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATTCA 2608
QY 901 GCTCATGTTAACTCACTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCAG 960
Db 2609 GCTCATGTTAACTCACTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCAG 2668
QY 961 TCTGTTAGTGAATTAACCTATCAGCTTCAATCAACTTACCTTATGAGTCTTTAAAGCT 1020
Db 2669 TCTGTTAGTGAATTAACCTATCAGCTTCAATCAACTTACCTTATGAGTCTTTAAAGCT 2728
QY 1021 ATTTAAAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCATTCACCCCTAACA 1080
Db 2729 ATTTAAAGTGAACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCATTCACCCCTAACA 2788
QY 1081 CCAGCGGATGAGGAAGAGTTTATCAATTAATTAACACAGTACGACCAAAAC 1134
Db 2789 CCAGCGGATGAGGAAGAGTTTATCAATTAATTAACACAGTACGACCAAAAC 2842
RESULT 6
AB014136 51896 bp DNA linear BCT 19-JUL-2002
LOCUS Streptococcus pyogenes MGAS315, section 1 of 37 of the complete
DEFINITION genome.
ACCESSION AB014136 AE014074
VERSION AE014136.1 GI:21903716
KEYWORDS
SOURCE Streptococcus pyogenes MGAS315
ORGANISM Streptococcus pyogenes MGAS315
Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 51896)
AUTHORS Beres,S.B., Sylva,G.L., Barbican,K.D., Lei,B., Hoff,J.S.,
Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F.,
Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and
Musser,J.M.
TITLE Genome sequence of a serotype M3 strain of group A Streptococcus:
Phage-encoded toxins, the high-virulence phenotype, and clone
emergence
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10078-10083 (2002)
REFERENCE 2 (bases 1 to 51896)
AUTHORS Beres,S.B., Sylva,G.L., Barbican,K.D., Lei,B., Hoff,J.S.,
Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F.,
Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and
Musser,J.M.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,
Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,
Hamilton, MT 59840, USA
FEATURES
source Location/Qualifiers
1 .51896
/organism="Streptococcus pyogenes MGAS315"
/mol_type="genomic DNA"
/strain="MGAS315"
/serotype="M3"
/db_xref="taxon:198456"
/note="group: A"
232. .1587
/gene="dnaA"
/note="synonym: Spym3_0001"
232. .1587
/gene="dnaA"
/note="best blastp hit: gb|AAF71535.1|AF255728.1
(AF255728) chromosomal initiator protein DnaA
[Streptococcus pyogenes]"
/codon_start=1
/transl_table=11

/product="chromosomal initiator protein"
/protein_id="AAM78608.1"
/db_xref="GI:21903717"
/translation="MTENEQIFMNRVLELAQSQLKQATYEFFVHDARLLKVDKHIATI
YLDOMKELFWENKLDVLTAGFVYNAQISVDVVEEDLMIEQNOTKINOKPOCAL
NSLPTVSDLSKYSFENFIQDENRWAAASIAVANTPGTTPYPLFWGPGGLKTH
LNAIGSVLLENPNARIKYITAENFNEFVHRLDTMDELKFRNLDDLLLDIDQ
SLAKTUSGTQEEFFFNALHNNKQIVLSDRTPDHLNDELKRLVTRFKWGLTVNI
TPPDEFVAILTKIKIYENFIFFPDITIEYLAGQFDSNVRLEGLAKDISLVANFKQI
DTITVDAEAIARAKODGPRWTPIEBEIQAGVKFYGVTKETIKRTQNIIVLAR
QVAMFLAREMTDNLPLKIGKEFGCRDHSVLHAYNKIKNMI SQDESRIETIKNKI
K"
1742. .2878
/gene="dnaA"
/note="synonym: Spym3_0002"
1742. .2878
/gene="dnaA"
/note="best blastp hit: gb|AAK33147.1| (AE006472) beta
subunit of DNA polymerase III [Streptococcus pyogenes M1
GAS]"
/codon_start=1
/transl_table=11
/product="beta subunit of DNA polymerase III"
/protein_id="AAM78609.1"
/db_xref="GI:21903718"
/translation="MIQFSINRTLFIHALNATKRAISTKNAIPILSSIKIEVTPGT
LTGSGQISIENTIPVSNENAGLLTSPGAILLEASFIIISLPDININVEIQH
QVLTSGKSEITLKQDVQYPRQVSTENPLIKTKLKSIIAETAFASLQBSRP
ILTGHVILSNHKDFKAVATDSHRMSORLITLNTSADFVVI PKSLREFSAVFTDD
IETVEFSPSQILFRSEHISFYTRLLGNYPDTRLMTTEFETEVVENTOSLRHAME
RAFLSNATQNGTVKLEITQNHISAHVNSPEVGKNEDLDIVSGSGLTISFNPTYL
IESLKAIKSETVKIHLFSPVRPFTLTGDBEESFIQLITPVRTN"
2953. .3150
/gene="Spym3_0003"
2953. .3150
/gene="Spym3_0003"
/note="best non-GAS blastp hit: gb|AAC45338.1| (AF000658)
ORFX [Streptococcus pneumoniae]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM78610.1"
/db_xref="GI:21903719"
/translation="MYQIGSPVEMKKPHACVIKETGKKANQWVLVRGADIKIQCTNC
QHVIMSGRYDFERKLKKVLP"
3480. .4595
/gene="Spym3_0004"
3480. .4595
/gene="Spym3_0004"
/note="best blastp hit: gb|AAK33149.1| (AE006472) putative
GTP-binding protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="AAM78611.1"
/db_xref="GI:21903720"
/translation="MALTAGIVLGNVKGSTLFNAITKAGAAANYPPATIDPNVGMV
EVPDERLQKLTTELITPKKTVPTTFDITAGIVKAGSGEGLGNKFLANIREVDAIVH
VVRAPDDENVNREQRDAFVDVPVADITINLELILADLESINKRYARVEKMKATQKD
KESVAEENVLQKIKPVEDGKSARTIETDEAKVKGLFLLTTKPVLYVANVEDKDV
ANPDGIDYQKIRDFATENAENVVISARAEELLEDDEDEKEFLEAIGLTESGVDK
LTRAAYHLGLGTYFTAGEKEVRAVTFKGIKAPQAGIILHSDFERGFIRAVTMSYDD
LMTYSSEKAVKEAGRLREEGKBYVQDQDIMEFRFNV"
4665. .5234
/gene="pth"
/note="synonym: Spym3_0005"
4665. .5234
/gene="pth"
/note="best blastp hit: gb|AAK33150.1| (AE006472) putative
peptidyl-tRNA hydrolase [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative peptidyl-tRNA hydrolase"

```
/protein_id="AA078612.1"
/db_xref="GI:21903721"
/translation="MYKMTVGLNGSGKVEKTKNIGMAIDNIVKNDVFTTDDKRF
KQIGSTFINHEKVCEVKTTFNMNSGIAVXALTYINDITDLIVYDIDMEVSKL
RLRSKGSAGHNGIKSIIAHIGTQEFNRIKYGRPLKGMTVINHVNGQFNTEDNIAI
SUTLDRVNAVAFYLOENDFEKTMQKFG"
5321..8740
/gene="trcf"
/notes="synonym: Spym3_0006"
5321..8740
/gene="trcf"
/notes="best blastp hit: gb|AAK33151.1| (AE006472) putative
transcription-repair coupling factor [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative transcription-repair coupling factor"
/protein_id="AA078613.1"
/db_xref="GI:21903722"
/translation="MGLSGSSKALAIASAYLDQKKIVVVTSTONEVEKLADLSLL
DELVFFADVAABEFIFASMDKALSRITQLFLRNPKSQGLVLSLGLRTLLPN
PVTFTKSIQLTVGEDVSDTLTKMLTIGYQKVSQVISPGESESGDILDIEYIQE
LPYLEFGDDIDSIRQHPETQKSPOLSGVINPASDLIFASDFQKISQLEKAL
QTAQDKKSYLEDVLAISKNGFKHDIRKQSLFYKESWLLDIIPKGPPIFFDFQK
LVDKNARFLEIANLPDILQCKKALSNLNFADNRYELRHYPATFFNFHKLGNL
KFDNMHOLTOYANQEFNPFLLIDEIKRYQKQNTTVIQQVESQYAYERLEKSFQDQ
FRPLVSAQNVISRESQIVIGASSGFYFADEKALITEHYHKKIKRRARRSNISN
AERLKDNEAVGSDYVHNVHNGRFLGIEITIQIQRHYDVTIOYNSDRISLSDIQ
ISSLKTVASDEKPKINLNGRFQTKQKVARQVEDIADLLKLAERSQKQSF
SPDDLQRAFPDPAFVETEDLSRIKEIKADMESMOPMDRLLVGDVGFQKTEVAMRA
AFKAVNDKQVAVLPVTVLAQQHYENFKARPNFVENVLSRFSKKEQATLERY
PIRTHLMSGLRDLISVIEPTPNRPVQTVLENNPGLVREALIREMDRGQIFVY
YKVDIETKVAELSGPEASIFGVHQMSEIOLENTLIDFINGDYDLVATTIET
GVDSINVTLIFENADHMLSTLYQLRGVRSNRISAVAYLRPDKVLTVESKRL
AIGFTELSGFKAMRDLISIRGAGNLCASQSGFDSVGFPMYSQLEQAIASKQK
TTPYRGNTENIQLDAYLPDDYIADROKIDIKYRIEIOSREDYLNQDELIDRFG
EYPDVAYLLEIALLKHYMDNAFAELVEKKNQVIVPEVLSLSFLQDYFEALSKT
HLKAKISEHQKIDIVFDVRHQDKVRIILEELMLFGERLSEIKIRKNSVFK"
8902..9174
/gene="Spym3_0007"
8902..9174
/gene="Spym3_0007"
/notes="best non-GAS blastp hit: gb|AAK74200.1| (AE007318)
S4 domain protein [Streptococcus pneumoniae TIGR4]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA078614.1"
/db_xref="GI:21903723"
/translation="MRLDKYLKVSRLIKRRSVAKVADKGRIVKNGILAKSSTNKLN
DHIEISFGNKLGLTVRVIEIKDSTKEDALKWYELISETRITLNEA"
9161..9532
/gene="divic"
/notes="synonym: Spym3_0008"
9161..9532
/gene="divic"
/notes="best blastp hit: gb|AAK33153.1| (AE006472) putative
cell division protein (Divic) [Streptococcus pyogenes M1
GAS]"
/codon_start=1
/transl_table=11
/product="putative cell division protein"
/protein_id="AA078615.1"
/db_xref="GI:21903724"
/translation="MKKPSIVQLNNHYIKKENLKKXFEESQKRNFRMGWILVSNMF
LFILPTYNLVKSYVDPEKQVVKLKEYNELSKTKKEQLAERLKDNDNFVKIYAR
9667..10953
/gene="Spym3_0009"
9667..10953
/gene="Spym3_0009"
/notes="best non-GAS blastp hit: gb|AAK98813.1| (AE008385)
```

```
Conserved hypothetical protein [Streptococcus pneumoniae
R61]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA078616.1"
/db_xref="GI:21903725"
/translation="MRKULAAWLMFTFPLPVIKSTKLFISKNAVYQLKQVQVST

Query Match      98.3%; Score 1114.8; DB 1; Length 51896;
Best Local Similarity 98.9%; Pred. No. 2.2e-195;
Matches 1122; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      1  ATGATTCAATTTTCAATTAATCGCACATATTATTATTCATGCTTTTAAATACAACATAACGT 60
      DB      1742  ATGATTCAATTTTCAATTAATCGCACATATTATTATTCATGCTTTTAAATACAACATAACGT 60
      QY      61  GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAATTTGAAGTCACATCT 120
      DB      1802  GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAATTTGAAGTCACATCT 120
      QY      121  ACAGGAGTAACCTTTAAACAGGCTCTAACGGTCAATATCAATTTGAAACACTATTTCCTGTA 180
      DB      1862  ACAGGAGTAACCTTTAAACAGGCTCTAACGGTCAATATCAATTTGAAACACTATTTCCTGTA 180
      QY      181  AGTAATGAAATGCTGCTTTGCTAAATACCTCTCCAGGACTATTTTATTAGAGTACT 240
      DB      1922  AGTAATGAAATGCTGCTTTGCTAAATACCTCTCCAGGACTATTTTATTAGAGTACT 1981
      QY      241  TTTTATTATTAATTTTCAAGTTTGCAGATATTAGTATAAATTTTAAAGAAATTTGAA 300
      DB      1982  TTTTATTATTAATTTTCAAGTTTGCAGATATTAGTATAAATTTTAAAGAAATTTGAA 2041
      QY      301  CAACACCAAGTTGTTTTAACCACTGTGTAATCAGAGATTACCTTTAAAGAAAGAAAGATGT 360
      DB      2042  CAACACCAAGTTGTTTTAACCACTGTGTAATCAGAGATTACCTTTAAAGAAAGAAAGATGT 2101
      QY      361  GACCAGTATCTCTGTCTACAGAGATCAACAGAAATCCTTTGATTTTAAAAACAAAA 420
      DB      2102  GACCAGTATCTCTGTCTACAGAGATCAACAGAAATCCTTTGATTTTAAAAACAAAA 2161
      QY      421  TTATCAAGTCTATTATTCTGCAAAACAGCTTTTGCAGCCAGTTTACAGAGAGTCTCT 480
      DB      2162  TTATCAAGTCTATTATTCTGCAAAACAGCTTTTGCAGCCAGTTTACAGAGAGTCTCT 2221
      QY      481  ATTTTAAACAGGAGTTCATATTTAGTAATCAATAAGATTTTAAAGCAGTAGCGACT 540
      DB      2222  ATTTTAAACAGGAGTTCATATTTAGTAATCAATAAGATTTTAAAGCAGTAGCGACT 2281
      QY      541  GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGACAATCTTACAGAGATTGATG 600
      DB      2282  GACTCTCATCGTATGAGCAACGTTTAAATCACTTTGGACAATCTTACAGAGATTGATG 2341
      QY      601  GTAGTCTCTCCAGTAATCTTTGAGAGATTTTTCAGAGTATTTCAGAGATGATATTGAG 660
      DB      2342  GTAGTCTCTCCAGTAATCTTTGAGAGATTTTTCAGAGTATTTCAGAGATGATATTGAG 2401
      QY      661  ACCGTGTAGGTAATTTTCTCACCAGCAATCTTTGAGAGTGAACACATTTCTTTT 720
      DB      2402  ACCGTGTAGGTAATTTTCTCACCAGCAATCTTTGAGAGTGAACACATTTCTTTT 2461
      QY      721  TATACAGCCTCTTAGAGGAAATTTCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
      DB      2462  TATACAGCCTCTTAGAGGAAATTTCCCGATACAGACCGTTTATTAATGACAGAAATTT 2521
      QY      781  GAGACGGAGGTGTGTTTTCAATACCAATCCCTTCGCCACGCTATGAAACGTGCTCTTG 840
      DB      2522  GAGACGGAGGTGTGTTTTCAATACCAATCCCTTCGCCACGCTATGAAACGTGCTCTTG 2581
      QY      841  ATTTCTAATGCTACTCAAAATGCTACTGTGTTAAGCTTGAAGTACTCAAAATCATATTTCA 900
      DB      2582  ATTTCTAATGCTACTCAAAATGCTACTGTGTTAAGCTTGAAGTACTCAAAATCATATTTCA 2641
```

QY 901 GCTCATGTTAACTCACCCTGAGGTCGTAAGGTAACGAGGATTAGATATTGTTAGTCAG 960
|||||
Db 2642 GCTCATGTTAACTCACCCTGAGGTCGTAAGGTAACGAGGATTAGATATTGTTAGTCAG 2701
|||||
QY 961 TCTGCTAGTGAATTAACATATCATGCTTCAATCAACTTACCTTATTGAGTCTTTAAAGCT 1020
|||||
Db 2702 TCTGCTAGTGAATTAACATATCATGCTTCAATCAACTTACCTTATTGAGTCTTTAAAGCT 2761
|||||
QY 1021 ATTAAGAAGTGAACAGTAAATTCATTCTTATCACCAGTTCGACCATTCACCTTAACA 1080
|||||
Db 2762 ATTAAGAAGTGAACAGTAAATTCATTCTTATCACCAGTTCGACCATTCACCTTAACA 2821
|||||
QY 1081 CCAGCGCATGAGGAAGAAGTTTTCATCAATTAATTAACACAGTACGACCAAC 1134
|||||
Db 2822 CCAGCGCATGAGGAAGAAGTTTTCATCAATTAATTAACACAGTACGACCAAC 2875
|||||

RESULT 7
AP005141
LOCUS AP005141 311600 bp DNA linear BCT 07-APR-2004
DEFINITION Streptococcus pyogenes SSI-1 DNA, complete genome, section 1/6.
ACCESSION AP005141 BA000034
VERSION AP005141.1 GI:28810157
KEYWORDS Streptococcus pyogenes SSI-1
SOURCE Streptococcus pyogenes SSI-1
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

1 Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y.,
Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T.,
Hayashi, H., Hattori, M. and Hamada, S.
TITLE Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
large-scale genomic rearrangement in invasive strains and new
insights into phage evolution
JOURNAL Genome Res. 13 (6A), 1042-1055 (2003)
MEDLINE 22683278
PubMed 12799345

REFERENCE 2 (bases 1 to 311600)
AUTHORS Hattori, M., Yamashita, A., Oshima, K. and Shiba, T.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) Masahira Hattori, Kitasato Institute for
Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamiara,
Kanagawa 228-8555, Japan
(E-mail: hattori@genome.is.kitasato-u.ac.jp;
URL: http://genome.is.kitasato-u.ac.jp/; Tel: 81-42-778-8194,
Fax: 81-42-778-8193)

COMMENT This work was done in collaboration with Ichiro Nakagawa, Ken
Kurokawa, Masanobu Nakata, Yusuke Tomiyasu, Nobuo Okahashi,
Shigetada Kawabata, Kiyoshi Yamazaki, Teruo Yasunaga, Shigeyuki
Hamada (Osaka University), Hideo Hayashi (University of Tsukuba),
and supported by the Research for the Future Program of the Japan
Society for the Promotion of Science. This clone was isolated from
a patient presenting with toxic shock like syndrome.

FEATURES
source
1. 311600
/organism="Streptococcus pyogenes SSI-1"
/mol_type="genomic DNA"
/strain="SSI-1"
/db_xref="taxon:193567"
232..1587
/gene="SPS0001"
232..1587
/gene="SPS0001"
/note="similar to GB:AAF71535.1 (AP255728) percent
identity 100 in 451 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="chromosomal initiator protein"
/protein_id="BAC63096.1"
/db_xref="GI:28810158"
/translation="MTENEQIFWNRVLEAQSOLKQATYEFVVDARLLKVDKHIATY

YLDQKELFWKKNLKDVLITAGFEVYNAQISVDYVFEEDLMIEQNQTKINQPKQOAL
NSLPTVSDLSNYSFNIQGDENRWAAASIAVANTPTGTTIPLFIWGGGLGKTH
LLNAIGNSVLLENFARIKYITAEFNIFEVHILRDTMDLKEKRLVTRFKWGLTVNI
SLAKTSLGTOEEFNTFNHNNKQIVLTSRTPDHLNDEKRLVTRFKWGLTVNI
TPPFETRVAILTNKINFNALFIQDITTEYLAGQFDSNVRDLEGALKDISLVANFKOI
DTITVDIAABAIARAKODGPKMTVIPTEIEIOAQVKFYGVTKETKATKRTQNIVLAR
QVAMFLAREMTDNSLPKIGKEFGGRDHSVTLHAYNKIKNMI SQDESLEIEIETIKNKI
K"
1742..2878
/gene="SPS0002"
1742..2878
/gene="SPS0002"
/note="similar to GB:AAL96838.1 (AE009953) percent
identity 99 in 378 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative DNA polymerase III beta subunit"
/protein_id="BAC63097.1"
/db_xref="GI:28810159"
/translation="MIQFSINRGLTIFAHALNATKRAISTKNAIPILSSIKIEVPTGVT
LTGNGQISIENTIPVSNAGLLITSPGAILLEASFNIISSLPDININVKLEQH
QVLTSGKSEITLKGKVDQYPRLOEVSTENPLIKTLKLSIIAETAPASLOSRRP
ILTGHVILSNHDKFKAATDSHMSORLITLNTSADFVDVVIKGSIRESAVTD
IETVVFSPSQILFRSEHISFYTRLLEGNYPDTRLLMTFEFEVWVTSQSLRAME
RAFLISNATQNGTIVKLEITQNHISAHVNSPEVKVNDLIDIVSQSGDLTISFNPTYL
IESLIKAIKSETVKIHLSPVRPFTLTGDEESFIQLITPVRTN"
2953..3150
/gene="SPS0003"
2953..3150
/gene="SPS0003"
/note="similar to GB:AAK33148.1 (AE006472) percent
identity 100 in 65 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC63098.1"
/db_xref="GI:28810160"
/translation="MYQIGSFVEMKPKHACVIKGTGKANQKVLRVGADIKIQCTNC
QHVIMGRYDPERKLVKVLQP"
3480..4595
/gene="SPS0004"
3480..4595
/gene="SPS0004"
/note="similar to GB:AAK33149.1 (AE006472) percent
identity 98 in 371 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="BAC63099.1"
/db_xref="GI:28810161"
/translation="MALTAGIVGLPNVSKSTLENAITKAGAAANYPEATIDPNVGMV
EVPDERLQKLTETLPKTVPTTFTTFTDIAGIVKASRGEGIKNKFARIEKRVATVH
VVRAPDDNNVREGEDAFVDPVADITINLELILADLESINKRYKARIEKRVATVH
KESVAFNVLQIKPVLGDSKARTIETDEEAKVKGLFLITTKPVLVYRANVSDKV
ANPDGIDYVKQIRDFAAENAEVVIISARAEELIETDEEKEFLEAIGTESGVDK
LMTAAYHLGLGTYFTAGEKVRATFKRGIKAPQACGI IHSDFERGFIKRVATMSYDD
LMTYSSEKAVKEAGRLREEGKEYVVQDGDINEFRNV"
4665..5234
/gene="SPS0005"
4665..5234
/gene="SPS0005"
/note="similar to GB:AAK33150.1 (AE006472) percent
identity 99 in 189 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative peptidyl-tRNA hydrolase"
/protein_id="BAC63100.1"
/db_xref="GI:28810162"
/translation="MVKMI VGLGNPGSKYEKTKINIGPMATDNI VKNLDVITFTDDKNF


```
KAQIGSTFINHEKVCVFVKPTTFMNNNSGIAVKALLTYNNIDITDLIVYDDLDMEVSKL
RLRSKSGAGHNGIKSIIAHIGTQBFNRIKGRPLKGMVINHWGQFNTDNIAI
SUTLDRVNVAKFYLQENDFEKTMQKENG"
/gene="SPS0006"
CDS
5237..8740
/gene="SPS0006"
/note="similar to GB:AAK33151.1 (AE006472) percent
identity 99 in 1167 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative transcription-repair coupling factor"
/protein_id="BAC63101.1"
/db_xref="GI:28810163"
/translation="MDILELFSQWKVQSWHSGLTTLGROLVNLGSSSKALAIASAY
LDDQKIIVVSTQNEVKLASDLSLLDEELVQFPADDDVAAAEIPIASMKALSGRI
ETLOFLRNPKSQGLVIVLSGLRLTLNPNPVTFKSQILQTVGSDYSDITLTKQLMTIG
YKQVSQVISPGEFERRGDIIDIVETQELPYRLEFFGDDIDSIRQHPETOKSFEQLE
GVFNPSADLIFENASDRQRIEQLKALQADDDKSYLEDVLAVSGFKHDKIRKF
OSLYEKEWLSLDYIPKGTPIFFDDFKQVDRNARFDEIANILLTDLQCKALSNLN
YFADNYRELRLHPATPFPSNFKLGNIKIFQKHQLTQVAMOEFFNQPLLIDELKRY
QKQNTVIVQESQYAYERLEKSFQDYQFRLPLVSAOIVRSQIIVIGAISSGYPFA
DEKALITEHEIYHKIKRRARSNIISNAERLKDYNELAVGVVHNVHVGIRFLGIE
TIQGIHRDYVITIQVNSDRIISLPIDQISLSKYVSADGKPKINKLNDGRFQKTKQ
KVARVEDIADLLKLYAESQKGSFSDDDLQAFDDDFAFVETEDQLRSIKEIK
ADMESQPMRDLVGVDFGKTEVAMRAAFKAVNDHKQVAVLPTTVAQHYENFKA
REFNYPVDVLSRFKSKKEQATLELVRKQDIIIGHRLLSKDVVPSDGLGLVID
EORFGVKHETLKEIKTKVDLTATPIPTLHMSMIGIRLSLVETPPNRPVPO
TYLENNPGLVREAIIREMDRGQIIFYVYVNVKDIETKVAELQELVPRASIGFVHGQ
SEIQLNTLIDFNGDYDLVATVITETGVDISNVTNLFENADHMGSLTYQLRGVR
GRSNRIAYAYLMYRPDKVLTEVSEKLEAIKGFTELGSFKIAMRDLISIRGANTLGA
SQSGFIDSQVGFMYSQLLEQALASQKGTIVROKGNTEINQIDAYLPDXYADERK
IDIVKIRIQRISREDYINLQDELIDRFGBYPQVAYLLEALLKHYMDNAFAELVERK
NNQIVRPFVTSLSYLTQDYFEALSKTHLKAKISEHQGKIDIVDFVRHQDYRLIE
LMLFGRLESEIKIRKNNVSFK"
/gene="SPS0007"
CDS
8878..9174
/gene="SPS0007"
/note="similar to GB:AAK33152.1 (AE006472) percent
identity 100 in 98 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC63102.1"
/db_xref="GI:28810164"
/translation="MKRIVGEIMRLDKYLKSLRIKRSVAKEVADKGRIVKNGILAK
SSTNKLNDHHEISFGNKLTVRIEIKDSTKEDALKMVEIISETRITLNEBA"
3161..9532
/gene="SPS0008"
CDS
9161..9532
/gene="SPS0008"
/note="similar to GB:AAL96843.1 (AE009953) percent
identity 100 in 123 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative cell division protein (DivC)"
/protein_id="BAC63103.1"
/db_xref="GI:28810165"
/translation="MKKPSIVQLNNHYIKKENIKKKFEEESQKRNRFMGWILVSMVF
LFILPTYNLKSVDYDFKQKQVVKLKEYNELSKSTKKEQLAERLKDKNFVKYAR
AKYLSREGEMIIYPIGLLPK"
9529..9654
/gene="SPS0009"
CDS
9529..9654
/gene="SPS0009"
/note="unknown"
/codon_start=1
/evidence=not experimental
/transl_table=11
```

```
Query Match 98.3%; Score 1114.8; DB 1; Length 311600;
Best Local Similarity 98.9%; Pred. No. 1.5e-195;
Matches 1122; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGATTCAATTTTCAATTAATCGCACATTATTATTTCATGCTTTTAAATACAACTAAACGT 60
Db 1742 ATGATTCAATTTTCAATTAATCGCACATTATTATTTCATGCTTTTAAATACAACTAAACGT 1801
QY 61 GCTATTAGCAGCAAAAATGCCATTCCTATTCTTTTCATCAATAAAAATGAAAGTCATCT 120
Db 1802 GCTATTAGCAGCAAAAATGCCATTCCTATTCTTTTCATCAATAAAAATGAAAGTCATCT 1861
QY 121 ACAGAGTAACCTTTAAACAGGCTCTAACCGTCAATATCAATTAATCAATTAATTAAGAGTAA 180
Db 1862 ACAGAGTAACCTTTAAACAGGCTCTAACCGTCAATATCAATTAATTAATTAAGAGTAA 1921
QY 181 AGTAATGAAATGCTGGTTCCTTAATACCTCTCCAGAGCTATTATTATTAGAAGCTAGT 240
Db 1922 AGTAATGAAATGCTGGTTCCTTAATACCTCTCCAGAGCTATTATTATTAGAAGCTAGT 1981
QY 241 TTTTATTATAATTATTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATGAA 300
Db 1982 TTTTATTATAATTATTCAAGTTTGCAGATATTAGTATAAATGTTAAAGAAATGAA 2041
QY 301 CAACACCAAGTGTGTTTAAACAGTGGTAAATCGAGATTACCTTTAAAGGAAAGATGTT 360
Db 2042 CAACACCAAGTGTGTTTAAACAGTGGTAAATCGAGATTACCTTTAAAGGAAAGATGTT 2101
QY 361 GACCAGTATCCTCGCTACAAAGATCAACAGAAATCCTTTGATTTTAAAAACAAA 420
Db 2102 GACCAGTATCCTCGCTACAAAGATCAACAGAAATCCTTTGATTTTAAAAACAAA 2161
QY 421 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCCTCT 480
Db 2162 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCCTCT 2221
QY 481 ATTTTAAACAGAGTTCATATTGTTTAAAGTAAATCATATAAGATTTTAAACAGTAGGACT 540
Db 2222 ATTTTAAACAGAGTTCATATTGTTTAAAGTAAATCATATAAGATTTTAAACAGTAGGACT 2281
QY 541 GACTCTCATGTAAGAGCAAGCTTTTAAATCACTTTGGACAATCTTCAGCAGATTGATG 600
Db 2282 GACTCTCATGTAAGAGCAAGCTTTTAAATCACTTTGGACAATCTTCAGCAGATTGATG 2341
QY 601 GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGCAGATTATTTACAGATGATTTGAG 660
Db 2342 GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGCAGATTATTTACAGATGATTTGAG 2401
QY 661 ACCGTTGAGGATTTTCTCAACCAAGCCAAATCTTTGTTTCAAGAGTGAACACATTTCTTTT 720
Db 2402 ACCGTTGAGGATTTTCTCAACCAAGCCAAATCTTTGTTTCAAGAGTGAACACATTTCTTTT 2461
QY 721 TATACAGGCTCTTGAAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
Db 2462 TATACAGGCTCTTGAAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 2521
QY 781 GAGACGAGGTTGTTTTCATACCAATCCCTTCCGACGCTATGGAACGTCCTCTTTG 840
Db 2522 GAGACGAGGTTGTTTTCATACCAATCCCTTCCGACGCTATGGAACGTCCTCTTTG 2581
QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAATCATATTTC 900
Db 2582 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAATCATATTTC 2641
QY 901 GCTCATGTTAACTCACCTGAGGTTGTAAGGTAAACGAGGATTTAGATATTGTTAGTCAG 960
Db 2642 GCTCATGTTAACTCACCTGAGGTTGTAAGGTAAACGAGGATTTAGATATTGTTAGTCAG 2701
QY 961 TCTGCTAGTGAATTTAACTCATCAGCTTCAATCCCACTTACCTTATGAGTCTTTAAAAAGCT 1020
Db 2702 TCTGCTAGTGAATTTAACTCATCAGCTTCAATCCCACTTACCTTATGAGTCTTTAAAAAGCT 2761
```



```
QY 1021 ATTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCTTAACA 1080
Db 2762 ATTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCTTAACA 2821
QY 1081 CCAGGCGATGAGGAGAAAGTTTATCCAAATTAATTAACACCAAGTACGACAAAC 1134
Db 2822 CCAGGCGATGAGGAGAAAGTTTATCCAAATTAATTAACACCAAGTACGACAAAC 2875

RESULT 8
LOCUS AR157860 3200 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 2 from patent US 6245906.
ACCESSION AR157860
VERSION AR157860.1 GI:16218975
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3200)
AUTHORS Ueyama,H., Abe,K., Keshi,H. and Matsuhisa,A.
TITLE Probes for the diagnosis of infections caused by Streptococcus
pyogenes
JOURNAL Patent: US 6245906-A 2 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..3200
/mol_type="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 74.9%; Score 849; DB 6; Length 3200;
Best Local Similarity 98.7%; Pred. No. 2.7e-146;
Matches 866; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 ATGATTCATTTTCAATTAATCGCACATTAATTCATGCTTTAAATACAACTAAACGT 60
Db 2324 ATGATTCATTTTCAATTAATCGCACATTAATTCATGCTTTAAATACAACTAAACGT 2383
QY 61 GCTATTAGCACTAAATAGCCATCTCTATCTTTCATCAATAAATGAAGTCACTTCT 120
Db 2384 GCTATTAGCACTAAATAGCCATCTCTATCTTTCATCAATAAATGAAGTCACTTCT 2443
QY 121 ACAGAGTAACTTTAACAGGGTCTAACGGTCAATCAATCAATGAAACACTATTCCTGTA 180
Db 2444 ACAGAGTAACTTTAACAGGGTCTAACGGTCAATCAATCAATGAAACACTATTCCTGTA 2503
QY 181 AGTAATGAAATGCTGGTTTGCTAATTAACCTCTCCAGGAGCTATTTATTAGAAGTAGT 240
Db 2504 AGTAATGAAATGCTGGTTTGCTAATTAACCTCTCCAGGAGCTATTTATTAGAAGTAGT 2563
QY 241 TTTTATTAATAATTAATTTCAAGTTTGCAGATATTAGTATATAATGTTAAAGAAATGAA 300
Db 2564 TTTTATTAATAATTAATTTCAAGTTTGCAGATATTAGTATATAATGTTAAAGAAATGAA 2623
QY 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360
Db 2624 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 2683
QY 361 GACCAGTATCCTCGTCTACAAGAAGTATCAACAGAAATCCTTTGATTTTAAACCAAAA 420
Db 2684 GACCAGTATCCTCGTCTACAAGAAGTATCAACAGAAATCCTTTGATTTTAAACCAAAA 2743
QY 421 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCGAGCCAGTTTACAAGAAAGTCGTCCT 480
Db 2744 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCGAGCCAGTTTACAAGAAAGTCGTCCT 2803
QY 481 ATTTTAAACAGGAGTTTCATATTGTTATTAAGTATATCAATAAGATTTTAAAGCAGTAGGACT 540
Db 2804 ATTTTAAACAGGAGTTTCATATTGTTATTAAGCAATATCAATAAGATTTTAAAGCAGTAGGACT 2863
QY 541 GACTCTCATGTTATGAGCCAAAGTTTAAATCACTTTTGGAC-AATACTTCAGCAGATTGAT 599
Db 2864 GACTCTCATGTTATGAGCCAAAGTTTAAATCACTTTTGGACAAATATCTTCAGCAGATTGAT 2923
```

```
QY 600 GGTAGTCTTCCAAAGTAAATCTTTGAGAGAATTTTTCAGCAGTATTTTACAGATGATTTGA 659
Db 2924 TGTGTTTATCCAAAGTAAATCTTTGAGAGAATTTTTCAGCAGTATTTTACAGATGATTTGA 2983
QY 660 GACCGTTCAGGTATTTTCTCACCAGGCCAATCTTGTTCAGAAGTGAACACATTTCTTT 719
Db 2984 GACCGTTCAGGTATTTTCTCACCAGGCCAATCTTGTTCAGAAGTGAACACATTTCTTT 3043
QY 720 TTATACAGCCCTCTTAGAAGGAAATTTATCCCGATACAGACCGCTTTATTAAATGACAGAA 779
Db 3044 TTATACAGCCCTCTTAGAAGGAAATTTATCCCGATACAGATCGTTATTAAATGACAGAA 3103
QY 780 TGAGACGGAGGTGTTTTCATATACCAATCCCTCCGACAGCTATGGAACGTCCTTCTT 839
Db 3104 TGAGACGGAGGTGTTTTCATATACCAATCCCTCCGACAGCTATGGAACGTCCTTCTT 3163
QY 840 GATTTCTAATGCTACTCAAAATGGTACTGTTAAAGCTT 876
Db 3164 GATTTCTAATGCTACTCAAAATGGTACTGTTAAAGCTT 3200

RESULT 9
LOCUS E17380 3200 bp DNA linear PAT 28-JUL-1999
DEFINITION gDNA derived from streptococcus pyogenes.
ACCESSION E17380
VERSION E17380.1 GI:5712063
KEYWORDS JP 1998262698-A/2.
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 3200)
AUTHORS Kamiyama,H., Fukuda,K., Karashi,H. and Matsuhisa,A.
TITLE DIAGNOSTIC PROBE FOR INFECTIOUS DISEASE CAUSED BY STREPTOCOCCUS
PYOGENES MICROBE
JOURNAL Patent: JP 1998262698-A 2 06-OCT-1998;
COMMENT FUSO YAKUJIN KOGYO KK
OS Streptococcus pyogenes
PN JP 1998262698-A/2
PD 06-OCT-1998
PF 25-MAR-1997 JP 1997071077
PI KAWIYAMA HIROSHI, FUKUDA KANAKO, KARASHI HIROYUKI, PI
MATSUHIISA AKIO
PC C12Q1/68,C07H21/04//C12N15/09,(C12Q1/68,C12R1:46),(C12N15/09,
C12R1:46);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
FH Key Location/Qualifiers
FT source 1..3200
/strain='SP-7-44'.
/organism='Streptococcus pyogenes'
FT location/Qualifiers
1..3200
/organism='Streptococcus pyogenes'
/mol_type='genomic DNA'
/db_xref='taxon:1314'

FEATURES
source
ORIGIN
Query Match 74.9%; Score 849; DB 6; Length 3200;
Best Local Similarity 98.7%; Pred. No. 2.7e-146;
Matches 866; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 ATGATTCATTTTCAATTAATCGCACATTAATTCATGCTTTAAATACAACTAAACGT 60
Db 2324 ATGATTCATTTTCAATTAATCGCACATTAATTCATGCTTTAAATACAACTAAACGT 2383
QY 61 GCTATTAGCACTAAATAGCCATCTCTATCTTTCATCAATAAATGAAGTCACTTCT 120
Db 2384 GCTATTAGCACTAAATAGCCATCTCTATCTTTCATCAATAAATGAAGTCACTTCT 120
QY 121 ACAGAGTAACTTTAACAGGGTCTAACGGTCAATCAATCAATGAAACACTATTCCTGTA 180
Db 2444 ACAGAGTAACTTTAACAGGGTCTAACGGTCAATCAATCAATGAAACACTATTCCTGTA 2503
QY 181 AGTAATGAAATGCTGGTTTGCTAATTAACCTCTCCAGGAGCTATTTATTAGAAGTAGT 240
Db 2504 AGTAATGAAATGCTGGTTTGCTAATTAACCTCTCCAGGAGCTATTTATTAGAAGTAGT 2563
QY 241 TTTTATTAATAATTAATTTCAAGTTTGCAGATATTAGTATATAATGTTAAAGAAATGAA 300
Db 2564 TTTTATTAATAATTAATTTCAAGTTTGCAGATATTAGTATATAATGTTAAAGAAATGAA 2623
QY 301 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360
Db 2624 CAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 2683
QY 361 GACCAGTATCCTCGTCTACAAGAAGTATCAACAGAAATCCTTTGATTTTAAACCAAAA 420
Db 2684 GACCAGTATCCTCGTCTACAAGAAGTATCAACAGAAATCCTTTGATTTTAAACCAAAA 2743
QY 421 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCGAGCCAGTTTACAAGAAAGTCGTCCT 480
Db 2744 TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCGAGCCAGTTTACAAGAAAGTCGTCCT 2803
QY 481 ATTTTAAACAGGAGTTTCATATTGTTATTAAGTATATCAATAAGATTTTAAAGCAGTAGGACT 540
Db 2804 ATTTTAAACAGGAGTTTCATATTGTTATTAAGCAATATCAATAAGATTTTAAAGCAGTAGGACT 2863
QY 541 GACTCTCATGTTATGAGCCAAAGTTTAAATCACTTTTGGAC-AATACTTCAGCAGATTGAT 599
Db 2864 GACTCTCATGTTATGAGCCAAAGTTTAAATCACTTTTGGACAAATATCTTCAGCAGATTGAT 2923
```

QY 121 ACAGGAGTAACTTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACATATTCCTGTA 180
 Db 2444 ACAGGAGTAACTTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACATATTCCTGTA 2503
 QY 181 AGTAATGAAATGCTGGTTGCTAAATACCTCTCCAGGAGCTATTTATTAGAGCTAGT 240
 Db 2504 AGTAATGAAATGCTGGTTGCTAAATACCTCTCCAGGAGCTATTTATTAGAGCTAGT 2563
 QY 241 TTTTATTAAATATTTTCAAGTTTGCCAGATATTAGTATTAATGTTAAAGAAATGAA 300
 Db 2564 TTTTATTAAATATTTTCAAGTTTGCCAGATATTAGTATTAATGTTAAAGAAATGAA 2623
 QY 301 CAACACAGATGTTTAAACAGTGGTAATACAGATTTACCTTAAAGGAAAGATGTT 360
 Db 2624 CAACACAGATGTTTAAACAGTGGTAATACAGATTTACCTTAAAGGAAAGATGTT 2683
 QY 361 GACCAATATCCTGCTACAGAGATGATCAACAGAAATCCTTTGATTTTAAACAAA 420
 Db 2684 GACCAATATCCTGCTACAGAGATGATCAACAGAAATCCTTTGATTTTAAACAAA 2743
 QY 421 TTATTGAAGTCTATTATGCTGAAACAGCTTTTGCCAGCAGTTTACAAGAAAGTGTCT 480
 Db 2744 TTATTGAAGTCTATTATGCTGAAACAGCTTTTGCCAGCAGTTTACAAGAAAGTGTCT 2803
 QY 481 ATTTTAAAGGATTCATATGTTATTAAGTAATCAATAAAGATTTTAAAGCAGTAGCGACT 540
 Db 2804 ATTTTAAAGGATTCATATGTTATTAAGTAATCAATAAAGATTTTAAAGCAGTAGCGACT 2863
 QY 541 GACTCTCATGATGAGCAAGCTTTAATCACTTTGGAC-ATACTTCAGCAGATTGAT 599
 Db 2864 GACTCTCATGATGAGCAAGCTTTAATCACTTTGGACAAATCACTTCAGCAGATTGAT 2923
 QY 600 GGTAGTCTTCCAAGTAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATTTGA 659
 Db 2924 TGTGGTTATTCCAAGTAATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATTTGA 2983
 QY 660 GACCGTTAGGATTTTCTTCAAGCAAAATCTTTGTTTCAAGAGTGAACACATTTCTTT 719
 Db 2984 GACCGTTAGGATTTTCTTCAAGCAAAATCTTTGTTTCAAGAGTGAACACATTTCTTT 3043
 QY 720 TTATACACGCTCTTAGAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGATT 779
 Db 3044 TTATACACGCTCTTAGAGAGAAATTTATCCGATACAGATCGTTTATTAATGACAGATT 3103
 QY 780 TGAGACGAGGTGTTTTCATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 839
 Db 3104 TGAGACGAGGTGTTTTCATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTT 3163
 QY 840 GATTCTTAATGCTACTCAAAATGGTACTGTTAAGCTT 876
 Db 3164 GATTCTTAATGCTACTCAAAATGGTACTGTTAAGCTT 3200

RESULT 10
 LOCUS AE014853 14925 bp DNA linear BCT 25-OCT-2002
 DEFINITION Streptococcus mutans UA159 section 1 of 185 of the complete genome.
 ACCESSION AE014853 AE014133
 VERSION AE014853.1 GI:24376380
 KEYWORDS
 SOURCE Streptococcus mutans UA159
 ORGANISM Streptococcus mutans UA159
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 1 (bases 1 to 14925)
 AUTHORS Ajdic,D., McShan,W.M., McLaughlin,R.E., Savic,G., Chang,J.,
 Carson,M.B., Primeaux,C., Tian,R., Kenton,S., Jia,H., Lin,S.,
 Qian,Y., Li,S., Zhu,H., Najjar,F., Lai,H., White,J., Roe,B.A. and
 Ferretti,J.J.
 TITLE Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14434-14439 (2002)

22295063
 12397186
 2 (bases 1 to 14925)
 Ajdic,D., McShan,W.M., McLaughlin,R.E., Savic,G., Chang,J.,
 Carson,M.B., Primeaux,C., Tian,R., Kenton,S., Jia,H., Lin,S.,
 Qian,Y., Li,S., Zhu,H., Najjar,F., Lai,H., White,J., Roe,B.A. and
 Ferretti,J.J.
 Direct Submission
 Submitted (09-JUL-2002) Department of Microbiology and Immunology,
 University of Oklahoma Health Sciences Center, 940 SL Young Blvd.,
 Oklahoma City, OK 73104, USA
 Location/Qualifiers
 1. 14925
 /organism="Streptococcus mutans UA159"
 /mol_type="genomic DNA"
 /strain="UA159"
 /db_xref="taxon:210007"
 194. 1552
 /gene="dnaA"
 /locus_tag="SMU.01"
 194. 1552
 /gene="dnaA"
 /locus_tag="SMU.01"
 /note="Best Blastp Hit: gb|AAF71535.1|AF255728_1
 (AF255728) chromosomal initiator protein DnaA
 [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="chromosomal replication initiator protein, DnaA"
 /protein_id="AAN57794.1"
 /db_xref="GI:24376381"
 /translation="MTENEQIFWNRVLEAQSOLQKATDFFFVSDAKLLKVEGNIAT
 LLDMMKLFWKNIQPVLTAGFENVEIIEYVFEETQSTNSPOISQNKTAELAT
 ETLPFQVNDLPKSPFNFGVIGDENWAFSTASVADLPGTITNPLFIYGGPGUGKTH
 LNAIGNSVLASNPARKIYISAEFNFEFVHRLQMDLKKFNLDLLLLDDIQ
 SLAKSLAATQEEFFNTFALHNNKQIVLTSDTPDLDLEQLVTRFKWGLTVNI
 TPDEFTRVAILNNKIQYNSYSLSETIYLAGQFDSNVRDLECALDKISLVANFKKL
 DVIIVEAABAIRAKODSPFKMIVIPEDIQKQVGFYGVTVKEIKSTRTONIVLA
 ROVGYMLAREMTDNLSPKIGFGRDHSTVLHAYNKIKNLAQDDSLKIEMETIKNK
 IK"
 1708. 2844
 /gene="dnaA"
 /locus_tag="SMU.02"
 1708. 2844
 /gene="dnaA"
 /locus_tag="SMU.02"
 /EC_number="2.7.7.7"
 /note="Best Blastp Hit: sp|O06672|D3B_STRPN DNA
 POLYMERASE III, BETA CHAIN >gi|2109446|gb|AAC45337.1|
 (AF000658) beta subunit of DNA polymerase III
 [Streptococcus pneumoniae]"
 /codon_start=1
 /transl_table=11
 /product="putative DNA polymerase III, beta subunit"
 /protein_id="AAN57795.1"
 /db_xref="GI:24376382"
 /translation="MIKFSINKYFFIQLALNATRAISSKNAIPILSSLKTEVNSQSTT
 LTGNGQISTENTISAEENAGLLVTSSGAILLEAFNFINIVSSLPDITLDFEIEGH
 OVVLNGKSEITLKGKDVQYQPLQVGTNNPLILTKLTIITSETAFARQESRP
 ILGVHVLNTHKEFKAVATDSHRMSQKRLTDHSSDDFDVIVPSSLSRFAAFVTD
 IESVFFNSQILFRSEYISFYRLLEGVPTDRLGNFETVYFNTNLRAYME
 RAHLISNATQNGTKLEIINNQVTAHVNSPEVGVNEELDIESLGNDLTISNPPTVL
 IEALKAKSETVITRISPIRPTLTTPSDNSENFIQLITPVRTN"
 3106. 3297
 /locus_tag="SMU.05"
 3106. 3297
 /locus_tag="SMU.05"
 /note="Best Blastp Hit: gb|AAC45338.1| (AF000658) OREX
 [Streptococcus pneumoniae]"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAN57796.1"

/db_xref="GI:24376383"
/translation="MYELGSLVEMKKPHACTIKTTGKKANCWEVVRIGADIKIKCTNC
DRVNMSRHRDPERKIKVL"
3461..4576
/locus_tag="SMU.06"
3461..4576
/locus_tag="SMU.06"
/note="Best Blastp Hit: dbj|BAB07770.1| (AF001520)
GTP-binding protein [Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="AAN57797.1"
/db_xref="GI:24376384"
/translation="MALTAGIVGLPNVCKSLFNNAITKAGABAANYPPATIDPNVGMV
EVPDIRLQJLELIVPKKTVPTTEFTDIAGIVKSGEGELGNKFLANREIDAIVH
VPRFNDENVREQRESDVPDIADIETINLEILADLESINKRYARVEIKARTQK
KDSVAEFNLOKIKVLEDGKSARTIDFTEBDAKIVKGLFLLTKPVLVANVEDKV
ADPDSIDVQOIREPAATENAENVVISARAEESSELDEDRAPLEAMGLNESGVK
LTRAAYHLGLGTYFTAGEKEVRANTFKRGIKAKQAAGLIHSDFEKGFIRATWSYDD
LIKYSEKAVKAEAGLREEGKDYIVQDGDIMEFRNV"
4663..5232
/gene="pth"
/locus_tag="SMU.07"
4663..5232
/gene="pth"
/locus_tag="SMU.07"
/note="Best Blastp Hit: sp|P37470|PTH BACSU PEPTIDYL-TRNA
HYDROLASE (PTH) (STAGE V SPOULATION PROTEIN C)
>gi|2127242|pir||S66083 stage v sporulation protein -
Bacillus subtilis >gi|7448588|pir||C69715 stage v
sporulation protein spovc - Bacillus subtilis
>gi|467442|dbj|BAA05288.1| (D26185) stage v sporulation
[Bacillus subtilis] >gi|2632320|emb|CAB11829.1| (Z99104)
thermosensitive mutant blocks spore coat formation (stage
v sporulation) [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="putative peptidyl-trna hydrolase"
/protein_id="AAN57798.1"
/db_xref="GI:24376385"
/translation="MTKLIIVGLNPGSKYHETKHNVGPMALDRIVKPLNVSTENKTF
KAETVATPENGKIFYKPTFMNAGSLAVRAFLAYNLTVEDLIVYDLDMEVKI
RPRQKSGAGHNGIKSIKEIGTQBFRIKIGIRPKDKMTVNVVLGKFKDKBEITI
FNTLDKDKAVNYVLTQKTFQBTQKFGN"
5234..8722
/gene="trcf"
/locus_tag="SMU.08"
5234..8722
/gene="trcf"
/locus_tag="SMU.08"
/note="Best Blastp Hit: gb|AAD03810.1| (AF054624)
transcription-repair coupling factor [Lactobacillus
sakei]"
/codon_start=1
/transl_table=11
/product="putative transcription-repair coupling factor"
/protein_id="AAN57799.1"
/db_xref="GI:24376386"
/translation="MWLIDLFKQNKIKEWTEHLSAPTRQLVMGFSSSSKALATAAY
KQQKGIIVVSTQNEAREKLATDLISLIGEEVYQFPADVAAREFFIPASQDRNSRV
ESLNPFLNKSGLMSTIVGIRILLPSPYQAAQFKVGEENLDNIIKQLSNIG
YQKVAQVNPGEFSKRGDILDIYEINQPYRLEFPFGEIDGIRTFDDQDRSLNLD
SFILTPSDDIILITDSDFQACVNFLEKALITANDSLQPYLKEVLSVTQDHRQDVRFK
LSFFQKDWLLNLYLPEOTPLFFDFDQFIRIDNSHFDLIDIANLITEDLQNSKAVSSLS
YFANNYKDLRHQAPATFESNFKGLNLFKDYLYNFTQYPMQEPFNQPLIILDIRRY
QKSEMYLLQDSDQSQRLOKALQGNFNLPISQVDNLLKKQAQLTIGNLSTGFYFA
DEKVLITREIFHKVYKHKRRRSISNAERLKOYNELEKGDYVNVHVGIGRELFIE
TIEVSGVHRDYLTYQVADRISIPIEQIELLSKYVSDGKEPKINKLNDGRFQIKQ
KVSQVEDIADLLKLYAERSQLKGAFSPDQKNOEFPDFAVYETEDQLRSIKEVK
KDMEKQPMRLDVGDFGKTEVAMRAAFKAVNDKQVAVLPTVTLVLAQQHYNTFEK
RFENFPVNVAMMSRFKTEQSELTJLAKGQVDIIIGTHRLLSKQVTFKDLIVID
BEQRPGVHKETLKELTKVDVLTATPPIPRTHMSMLGIRDLSVISTPPSPNPVQ

TYUMETNASVIREAIMREIDRGQIFVIYNRVDTIDKKVAELRELIPEANTGFVHGKM
SEQUENTLDFISGVDLLVTTITTEIGDIDSNVNTLFIENADTMGLUSTLYQLRGRV
GRSRIAYALMHPDKILTEVSEKLEAIKGTGELSGFKIARMDLSIRAGNLTGS
SQSGFIDSQFEMYSQLEEAIAAQGRQKTRSKNAEINLIQDAYLPTNYSIDERQK
IYIKRIREISTKDYOALODELIDRFQDYPQVAYLIEIGLKAYLIDDAVESLDRK
QNSVVVPEKISLSKFLYTFDYFEALSKTNLKAISENKGKIEIVFDVRAKKEYEILEE
LIKFGKTLKKLKKKK"
8852..9124
/locus_tag="SMU.09"
8852..9124
/locus_tag="SMU.09"
/note="Best Blastp Hit: sp|P37557|YABO BACSU HYPOTHETICAL
9.7 KDA PROTEIN IN MFD-DIVIC INTERGENIC REGION
>gi|2127042|pir||S66089 conserved hypothetical protein
yabo - Bacillus subtilis >gi|467448|dbj|BAA05294.1|
>gi|2632326|emb|CAB11835.1| (Z99104) similar to
hypothetical proteins [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN57800.1"
/db_xref="GI:24376387"
/translation="WRLDKYLKVSRIIKRRSVAKEVSDKGRIKVGILAKSSDULKIN
DTIEIRFQNKLLTVRVLEMKDSTKKEDAVKMYEIISETRIDNGET"
9111..9482
/locus_tag="SMU.10"
9111..9482
/locus_tag="SMU.10"
/note="Best Blastp Hit: gb|AAC98903.1| (AF023181) Divic
homolog [Listeria monocytogenes]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN57801.1"
/db_xref="GI:24376388"
/translation="MEKPNIVQLNNKYNIDKTKRYEESTKRRHRFIMLILFIILL
Query Match 60.9%; Score 690.8; DB 1; Length 14925;
Best Local Similarity 75.6%; Pred. No. 2.3e-117;
Matches 857; Conservative 0; Mismatches 277; Indels 0; Gaps 0;
Qy 1 ATGATTCAATTTTCAATTAATCGCACATTATTATTATTCATGCTTTAAATACAACTAAAGT 60
Db 1708 ATGATAAAATTTTCAATTAATAAAGTTTTCCTTACAAAGCTTAAATGCTACCAAGCGA 1767
Qy 61 GCTATTAGCACTAAATGCGCATTCCTATTCTTTTCATCAATAAAATTTGAAGTCACTTCT 120
Db 1768 GCTATTAGTCTAAATGCTATTTCCTATTCTTTCTAGTTTAAATTTGAAGTGAATCT 1827
Qy 121 ACAGAGTAACTTTAAACAGGCTCTAAACGGTCAATATCAATATGAAAAACACTATTCTCTGA 180
Db 1828 CAGTCCATTACTTTAAACAGGCTCTAAATGGACAAATTTCTATTGAAAAATACGATTTTCA 1887
Qy 181 AGTAATGAAATGCTGGTTTGTCTAATTTACCTCTCCAGGAGCTATTTTATTAGAACTAGT 240
Db 1888 GAAGAAGAAATGCTGGACTATTGGTTACTTCTTTCAGGAGCTATCTTACTTTGAAGCTAAT 1947
Qy 241 TTTTATTATATATTATTTCAAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTTGAA 300
Db 1948 TTCTTTATATATTTGTTTCAAGCTACCTGATATCACTTTAGATTTTGAAGAAATTTGAA 2007
Qy 301 CAACACCAAGTTGTTTTAACAGTGGTAAATCAGAGATTACCTTTAAAAAGGAAAAAGATGTT 360
Db 2008 CAACATCAAGTTGTTTTAAATAGTGGAAAAATCAGAAAATACACTTTAAAGGTTAAAGATGTG 2067
Qy 361 GACCATATCTCGTCTACAGAAGATATCAACAGAAATCTTTGATTTTAAAAACAAAA 420
Db 2068 GAACAATATCTCGTTGTCAGNAAGTTGGAAACAAATAATCTCTTTAATTTTGAAGAAACAAA 2127
Qy 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGGCAGCAGTTTTCACAGAAAGTGGTCTCT 480
Db 2128 TTGTTAAAAACAATATTATTCTGAAACAGCTTTTTCGCTAGTACCAAGAAAGTGCACCA 2187

481	QY	ATTTTAA	CAGAGT	TTCA	TAT	TGT	TAT	TAA	GT	TAAT	CAT	AA	GA	AT	TTT	AA	GA	TC	AG	TAC	CG	ACT	540
2188	Db	ATTTTAA	CAGGT	TTCT	AT	TGG	TAT	TG	CA	AA	AT	CAT	AA	GA	TTT	AA	GA	TC	AG	TG	TG	CT	2247
541	QY	GACTCT	CAT	CGT	AT	G	AG	CA	CG	TTT	TA	T	CA	CT	TTT	G	CA	CA	T	ACT	CT	TG	600
2248	Db	GACTCT	CAT	CGAT	T	AG	T	CA	AG	AA	AT	T	AA	CT	CT	T	G	CA	CA	T	CT	TG	2307
601	QY	GTAGT	CTT	CCA	AG	TAA	AT	CTT	T	G	AG	AA	AT	TTT	T	G	CA	CA	G	T	AT	T	660
2308	Db	GTCTG	CAT	TCCA	AG	CCG	T	CT	TTT	A	CG	TG	AA	TTT	T	G	CA	CA	G	T	AT	T	2367
561	QY	ACGGT	T	CAG	TAT	TTT	T	CT	C	AA	G	CA	AA	AT	CT	T	G	T	T	C	AG	AG	720
2368	Db	TCGTG	T	CAG	GT	AT	TTT	T	CT	C	T	AA	C	AG	T	CA	AA	T	CT	T	T	AG	2427
721	QY	TATAC	CG	CT	CT	T	AG	AA	GA	AT	T	AT	CC	CA	G	AT	CA	CA	G	CG	TT	T	780
2428	Db	TATAC	GG	CC	CT	T	T	AG	AA	GA	AT	T	AT	CC	CA	G	AT	CA	CA	G	CG	TT	2487
781	QY	GAGAC	GG	AG	GT	T	G	TTT	CA	AT	T	CA	CC	CA	T	CG	CA	CG	CT	AT	T	G	840
2488	Db	GAACA	GA	AG	T	T	G	T	CT	CT	CA	AT	T	CA	T	AT	T	CA	T	AT	T	G	2547
841	QY	ATTTCT	TA	AT	G	CT	ACT	CA	AA	T	G	T	CT	GT	TA	AG	CT	T	C	AG	AT	T	900
2548	Db	ATCT	CA	AT	G	CT	AC	CA	AA	T	G	T	CT	GT	TA	AG	CT	T	C	AG	AT	T	2607
901	QY	GCT	CAT	GT	T	AA	CT	CAC	CT	G	AG	TT	T	G	GT	TA	AG	CT	TA	T	AT	T	960
2608	Db	GCAT	GT	T	AA	CT	CAC	CT	G	AG	TT	T	G	GT	TA	AG	CT	TA	T	AT	T	T	2667
961	QY	TC	T	G	T	AG	T	AG	T	CA	AT	T	CA	AT	T	CA	AT	T	CA	AT	T	CA	1020
2668	Db	TC	AG	AA	T	GA	TT	T	AA	CA	AT	T	AA	CA	AT	T	AA	CA	AT	T	AA	CA	2727
1021	QY	ATTTAA	AG	T	CA	AA	CAG	T	AAAA	TT	CA	TT	CT	T	TA	CT	C	CA	CG	TT	C	CA	1080
2728	Db	CIT	T	AA	AG	T	GA	CT	GT	C	AA	TT	CG	TT	T	AT	T	T	T	T	T	T	2787
1081	QY	CC	AG	CG	AT	G	AG	AA	AG	AT	T	T	AT	CC	AA	TT	AT	T	AT	CA	CG	T	1134
2788	Db	CC	AG	T	GA	AA	CAG	T	AAAA	CT	CA	TT	CA	AT	T	GA	T	T	CA	CT	GT	T	2841

```

RESULT 11
LOCUS      CQ648156
DEFINITION Sequence 5113 from Patent WO0234771.
ACCESSION  CQ648156
VERSION     CQ648156.1 GI:41693804
KEYWORDS
SOURCE      Streptococcus agalactiae
             Streptococcus agalactiae
             Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
             Streptococcus.
REFERENCE   1
AUTHORS     Telford,J., Masignani,V., Margarit y Ros,I., Grandi,G., Fraser,C.
             and Tettelin,H.
TITLE        Nucleic acids and proteins from streptococcus groups a & b
             Patent: WO 0234771-A 5113 02-MAY-2002;
             Chiron S.p.A. (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
            Location/Qualifiers
             1..1134
                /organism="Streptococcus agalactiae"
                /mol_type="unassigned DNA"
                /db_xref="taxon:1311"
ORIGIN

```

Query Match 56.5%; Score 641.2; DB 6; Length 1134;
Best Local Similarity 72.8%; Pred. No. 5.1e-108;
Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

QY	1	ATGATTCAAATTTTCAAATTAATCGCACATTAATTTATTTATTCATGCTTTAAATCAAACTAAACGTT	60
DB	1	ATGATTCAATTTTCAAATAAATAAAAAATTTCTTCTTGCACTCTAAACGGTAACCAACGA	60
QY	61	GCTATTAGCACTAAAAATGCCATTCCTATTTCTTTTCATCAATATAAAATTTGAAGTCACCTTCT	120
DB	61	GCTATTAGTCATAAAATATGCGATTTCCAAATCCTTTCAACTGTGTTAAATAGAAAGTGACTAGG	120
QY	121	ACAGGAGTAACCTTTAAACAGGGTCTAAACGGTCAAAATATCAATTTGAAAAACACTATTTCCTGTA	180
DB	121	GATGCTATCATTTTAAACGGGGTCAATTTGGAGCAAAATTTCAATTTGAAATATCTATTTCCTGCT	180
QY	181	AGTAATGAAATATGCTGGTTTGCTAAATTAACCTTCCAGGAGCTATTTTATTAGAACTAGT	240
DB	181	TCAAATGAAATATGACAGTTTACTAGTAACGAATCCAGGCTCTATTTGTTTAGAAGCTGGT	240
QY	241	TTTTTTTATTAATATTTATTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAAATTTGAA	300
DB	241	TTCTTTTATTAATATTTATTTCAAGTTTACAGATGTAATCTTTAGAAATTTACAGAGATTTGAA	300
QY	301	CAACACCAAGTTGTTTTAAACAGGTGGTAAATACAGAGATTACCTTTAAAGGAAAGATGTT	360
DB	301	CAACATCAAATTTGTTCTTACTAGTGGAAATACAGAGATTACTTTGAAAGGTAAAGGATGTC	360
QY	361	GACAGTATCCTCGTCTCAAGAGTATCAACAGAAATCCTTTGATTTTAAAAAACAAA	420
DB	361	GATCAATACCCCTCGTCTGCAGGAAATGACAAACAGATACCTCCATTAACATTTAGAAACTAAA	420
QY	421	TTATTGAAGTCTATTTATGTGTGAAACAGCTTTTTCAGCCAGTTTACAAAGAAAGTCGTCCT	480
DB	421	CTGTAAAAATCAATTTAATGAACTGCTTTTCTGCTAGCCAAACAGAAAGCCGTTCCA	480
QY	481	ATTTTAAACAGGATTCATATTTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT	540
DB	481	ATCTTAAACAGGTGTTCTATTTGGTTATCAGTCAAAATAAATACTTTAAGGCTGTTGCGACA	540
QY	541	GACTCTCATCGTATGAGCCAAAGTTTAAATCACTTTTGACAATACTTTCAGCAGATTTTGATG	600
DB	541	GATTCACACGATGAGTCAACGCACTTTCCAATTAGAGAAATCGGCTAATAATTTTGAT	600
QY	601	GTAGTTCTTCCAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATATTGAG	660
DB	601	TTGGTTGTTCCAAAGTAAATCCCTTCGAGAAATTTTCGGCTGTTTTTACAGATGATATTGAA	660
QY	661	ACGGTTCAGGTATTTTCTCACCAAGCCAAATCTTGTTCAGAGTGAACACATTTCTTTT	720
DB	661	ACTGTAGAGGTTTCTTCTCAGATAGTCAAAATGTTATTTAGAAGTGAATAATCAGCTTC	720
QY	721	TATACACGCCCTCTTAGAAGGAAATTTACCGATACAGACCGCTTTATTTAATGACAGAAATTT	780
DB	721	TATACACGTTTGGTTGAAGGAAACTACCTGTACTGATCGCCTCTCTAACTAATCAGTTT	780
QY	781	GAGACGAGGTTGTTTTCATACCCAAATCCCTTCGCCACGCTATGNAACGTGCTTCTTG	840
DB	781	GAAACCGAAATATCTTTTAATACAAATGCTTTTAGCCATGCTATGAAACGTGCTTATTTA	840
QY	841	ATTTCTAATGCTACTCAAAATGCTACTGTTTAAAGCTTGAGATTACTCAAAATCATATTTCA	900
DB	841	ATTTCGAATCGAACTCAGAACGCTACTGTTTCGTTTAGAAATTCAAATGAGACAGCTCA	900
QY	901	GCTCATGTTTAACTCACTGAGTTGGTAAAGTTAAACGAGATTTAGATATGTTTAGTCAAG	960
DB	901	GCTCATGTAAACTCTCCAGAAGTTGGTAAAGTTAAATGAGGAAATGGAATCTGTTAGTCTTT	960
QY	961	TCGTGTAGTGTATTTAATCATATCAGCTTCAATCCAACTTACCTTATTCAGTCTTTTAAAGCT	1020
DB	961	AAAGGTGATAGTTTAAATATATAGCTTTAATCCAACTTACCTAATGAACTTTTAAAGCA	1020
QY	1021	ATTTAAAGTGAAACAGTAAAAATTCATTTCTTATCACAGTTCGACATTCACCCCTAAACA	1080
DB	1021	GTAAAAAGCGAAACAGTTAGATTCGATTTATTTCTTCAGTAGCTCAATTTACTTTTGACA	1080

Qy	1081	CGAGCGATGAGGAAGAAGTTTATTCAATTAATTACACAGTACGAACAAC	1134
Dd	1081	CCTGGTGAAGATACTGAAGATTTTCATCCAATTAATAACTCCTGTCTGTAATAAC	1134
RESULT_12			
AX606582	AX606582	1137 bp	DNA linear PAT 17-FEB-2003
LOCUS	Sequence 4511 from Patent WO02092818.		
DEFINITION	AX606582		
ACCESSION	AX606582.1 GI:28402110		
VERSION			
KEYWORDS	Streptococcus agalactiae		
SOURCE	Streptococcus agalactiae		
ORGANISM	Streptococcus agalactiae		
REFERENCE	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
AUTHORS	Glaser,P., Rusniok,C., Chevalier,F., Frangeul,L., Lalioui,L., Zouine,M., Couve,E., Buchrieser,C., Poyart,C., Trieu-Cuong,P. and Kunst,F.		
TITLE	Streptococcus agalactiae genome sequence, use for developing vaccines, diagnostic tools, and for identifying therapeutic targets		
JOURNAL	Patent: WO 02092818-A 4511 21-NOV-2002;		
FEATURES	INSTITUT PASTEUR (PR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (PR)		
source	Location/Qualifiers		
ORIGIN	1..1137 /organism="Streptococcus agalactiae" /mol_type="unassigned DNA" /db_xref="taxon:1311"		
Query Match	56.5%;	Score 641.2;	DB 6; Length 1137;
Best Local Similarity	72.8%;	Pred. No. 5.1e-108;	
Matches	826;	Conservative 0;	Mismatches 308; Indels 0; Gaps 0;
Qy	1	ATGATTCAAATTTCAAATTAATCGCACATTAATTATTATCATGCTTTAAATACAACATAAAGT 60	
Dd	1	ATGATTCAATTTTCAATAAATAAAAAATTTCTTCTGCATGCTCTAACGGTAACCAACGA 60	
Qy	61	GCTATTAGCACTAAAATGCCATTCCTATTCTTTTCATCAATAAAAATTTGAAGTCATCTTCT 120	
Dd	61	GCTATTAGTCATAAAAATGGATTCCAAATCCTTTCAACTGTTAAATAGAAATGACTAGG 120	
Qy	121	ACAGGAGTAACCTTTAACAGGGTGCTAACCGGTCAAAATATCAATTGAAAAACACTATTTCCTGTA 180	
Dd	121	GATGCTATCAATTTNACGGGTCAATGGACAATTTCAATTGAAAATACATTTATCTCTGCT 180	
Qy	181	AGTAATGAAATGCTGGTTTGCTAAATTAACCTCTCCAGGAGCTATTTTATTAGAGCTAGT 240	
Dd	181	TCAAATGAAATGCAGGTTTACTAGTAGTAACGAATCCAGGCTCTATTTTGTAGAGCTGGT 240	
Qy	241	TTTTTTTAAATATATTTCNAGHTTTGCCAGATATTAGTATAATGTTTAAAGAAATTGAA 300	
Dd	241	TTCTTTTAAATATATTTCNAGFTTTACCAGATTAACAGATTAACCTTTAGAAATTTACAGAGATTGAA 300	
Qy	301	CAACACCAAGTTGTTTTAACCAAGTGGTAAATACAGAGATTACCTTTAAAGCAAAAGATGTT 360	
Dd	301	CAACATCAATTTGTTCTTACTAGTGGAAATACAGAGATTACTTTGAAAGGTAGGATGTC 360	
Qy	361	GACCAGTATCCTCGCTACAAGAAAGTATCAACAGAAAATCCTTTGATTTTAAAAAACAAA 420	
Dd	361	GATCAATACCCCTGCTGCAGGAAATGACAACAGATACTCCATTTAACATTTAGAAACTAAA 420	
Qy	421	TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCAGCCAGTTTACAAAGAAAGTCTGCTCT 480	
Dd	421	CTGTAAAAATCAATTTATTAAATGAACATGCTTTTCTGCTAGCCAAACAGAAAGCCGTCCA 480	
Qy	481	ATTTTAAACAGGAGTTTCATATTGTTATTAAAGTAATCATAAAGATTTTAAAGCAGATAGCGACT 540	
Dd	481	ATCTTAAACAGGTGTTCAATTTGGTTATCAGTCAAAATAAATATCTTTAAGGCTGTTTGCACA 540	
Qy	541	GACTCTCATCGTATGAGCCAAAGTTTAAATCACTTTTGACAATACTTCAGCAGAGATTTCGATG 600	

Scarselli, M., Mora, M., Iacobini, E.T., Brettoni, C., Galli, G.,
Mariani, M., Vegni, F., Maione, D., Rinaudo, D., Rappuoli, R.,
Telford, J.L., Kasper, D.L., Grandi, G. and Fraser, C.M.
Direct Submission
Submitted (18-JUL-2002) The Institute for Genomic Research, 9712
Medical Center Dr. Rockville, MD 20850, USA
FEATURES
Source
Location/Qualifiers
1. .21262
/organism="Streptococcus agalactiae 2603V/R"
/mol_type="genomic DNA"
/strain="2603V/R"
/serotype="v"
/db_xref="taxon:208435"
102. .1463
/gene="dnaA"
/note="synonym: SAG0001"
102. .1463
/gene="dnaA"
/note="identified by match to PFAM protein family HMM
PF00308"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein DnaA"
/protein_id="AAM98910.1"
/db_xref="GI:22533003"
/translation="MTENEQLFWRNVLSRSQIAAPAYEFFVLEARLLKIEHOTAVI
TLDNTEMKLFWQNLGVLTAPEIFNAEITANYVNSDLHLOETSPSNYQOSNEV
NLPFRKIDNLKERYTFANFVQGDENWASIASAVADSPGTTVPNLFVWGGPLGK
THLNAIGNVLRDNPANVLVITAENFINFVSHIRLDSMEELKEKEFNLDLLIDD
IQSLAKTLGGTGEFFENTFALHNDQVILTSDRNPQNLDELRVTPFSWGLPV
NITPDPFTRVALLTNKIQEYPDPDPIETFLQGFPSNVRELEGAKNISLVADFK
HAKTIVDIAEAIKARNKDGPIVTVPIEIEIQVGFYGVYVKEIKATKTDQIVL
ARQVAMYAREMNTDNLSPKIGKEFGGRDHTVLHAYNKIKNVAQDDNLRLEIETIKN
KIR"
1618. .2754
/gene="dnaN"
/note="synonym: SAG0002"
1618. .2754
/gene="dnaN"
/note="identified by match to PFAM protein family HMM
PF00712"
/codon_start=1
/transl_table=11
/product="DNA polymerase III, beta subunit"
/protein_id="AAM98910.1"
/db_xref="GI:22533004"
/translation="MIHFSINKNPFHLVTKRAISHKNAIPILSTVKIEVTRDAII
LTGNGQISIENTIPASNENAGLVNPGSILLBAGFEINIISSLPDVTLFETIEOH
QVLTSGKSEITLKGKVDQYPRLOEMTDTPLILEYKLLASINETAPASQOESRP
ILTGVLVISQNKYFKAVATSHRMSQRTPOLEKSNANFDLVVPSKSLRFSVAFTDD
IETVEFSDSCLFRSENISFYTRLLEGNYPDTRLITNQFETEIIFNTNALRHAME
RAYLISNATQNGTVRLSEINETSVAHNSPEVGNBELDVTSLKGSLSINISFNPTYL
IESLKAVKSETVTRFISPVRFPTLTGDEDTEDFIQLITPVRTN"
2824. .3705
/gene="SAG0003"
2824. .3705
/gene="SAG0003"
/note="identified by match to PFAM protein family HMM
PF00781"
/codon_start=1
/transl_table=11
/product="diacylglycerol kinase catalytic domain protein,
putative"
/protein_id="AAM98911.1"
/db_xref="GI:22533005"
/translation="MTLYLIANPHAGNKNASTIVGKIQELVHTEDISVFVTEOKDKRK
KQVINILSRPKESDHLMIIGDGTLSKVMYLPQHIPTVYPVPGSGNDPARKLGNL
KETITATQELKEINCFIVDKHLLNSLDLGFAYVWVASKSKIKNINLRKGLKI
TYIVIALSKLLHSKQVQLVEGETGQQIKLNDLYFFALANNYFGGGITITWPKASALT
AELDMVAKGHTFUKRLSILLSLVFKRHTTSKSKIKHTFKAMTVYFPKNSLIIDGSI
VELDQISLKKQKRLYM"
3715. .3912
/gene="SAG0004"
/note="identified by match to PFAM protein family HMM
PF00270"
/codon_start=1
/transl_table=11
/product="transcription-repair coupling factor"
/protein_id="AAM98916.1"
/db_xref="GI:22533010"
/translation="WNIEILFSQNKVVTRWHSGLVNTSNRQVLMGFGSKAKAIAASAY
EKLKIMVVTATQTDSDKLSLIGEDNVYQFADDDVPAAEIIFSLDLNKSIRL
SALRFLKDPKNGVLITSIGLRLLLPNPEVSKSQYKEIGQECVLDLCKNLNVLG
YKVSQVFSFGFSQRGDIDIFEMTQETPYRLEFFGDEIDGIRQFDIDTQKSLKQLE

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

SVQISPADDIILQADPERAKKLEGVLVTASEVQRTYLSVLSTTENHPKHSIRRP
LSIFYEKWGLDITPTEGPFVDDQKIVDRNAKLDLEIASLUTEDLQCKSHSSLN
YFSDPYKQRLQYQATPFSNPHKGLNLFKDLHFTQYGMQEFNPFLLVDINRY
KKSATVLLQVDSQKGLNLQENKEYGLDLIIISDKNDIVOKESQOLVHLSNGFYFA
DEKVLITIEREYHRRVKRRLRRSNIISNAERLKVKNLSVGVYVNVHVGKFLGIE
TIEQGLHRDVLTYQONADRSIPVEQIBELLTKVVSADGKEPKINTLNGRPFKAKQ
RVAKQVEDLADLLKLVAERSOLOGEAFSPDDNMQNDPNDPFAVETEDQLRSKEIK
QDMGNRPMDRLLVGDVGFGEKTEVEMARAFKAVNDHRQVVVLTIVLAQHPENFK
RFSNPVTVLVSFRPSKKEQDTLKRSLSGQVDIIIGTHRLLSQDVVFDLGLIVID
BEQFQVHKHEKJELKTKVDVLTATPIPTLHMSMLGIRLDSLVETPPTNRYPVQ
TYVLETPNGLVREAIIRIDRGGQVFFVYKNDVTDIDQVSELOELVPEASIGFVHGQ
SEIOLENTLIDFINGVDVLVATTIETGVDIISNVNTLFVENADHMLGLSTUYOLGRV
GRSRIAYVLMVYDPKVLTEISEKRLDAIKGFTLGSQPKIARDLSIRGAGNILGA
SQSGFISVGFEMYSQLEQAIATKQGLSRKQGNALQIDAYIPAEYISDEROK
IETIYRELETRADYALQDELIDRFEGEYDPQVAYLEIGLKKAYLIDLPAFTELVERK
GNEISIFEKASLKFITQDYFEALSKTQKARISETNGKMEVFNKHKKNYBIEE
LLKFAECFIEIKSRKPVEE"
9906..10001
/gene="SAG0009"
9906..10001
/gene="SAG0009"
/note="identified by Glimmer2; putative"
/codon_start=1

Query Match 56.5%; Score 641.2; DB 1; Length 21262;
Best Local Similarity 72.8%; Pred. No. 2.8e-108;
Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

Qy 1 ATGATTCATATTTCAATTAATCGCATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 1618 ATGATTCATATTTCAATTAATTAATAATTTCTTCTGATGCTCTAACCGTAAACCAACGA 1677
Qy 61 GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAAGTCACCTTCT 120
Db 1678 GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAAGTCACCTTCT 1737
Qy 121 ACAGAGTAACCTTAAACAGGCTCTAACCGTCAATATCAATTAAGAAACACTATTCTCTGTA 180
Db 1738 GATGCTATCATTTTAAACGGGGTCAAAATGGAACAAATTTCAATTAAGAAATCTATTCTCTGCT 1797
Qy 181 AGTAATGAAATGCTGGTTGCTAATACCTCTCCAGGAGCTATTTTATTAAGAAGCTAGT 240
Db 1798 TCAATGAAATGAGGTTTACTAGTACGAATCCAGGCTCTATTCTTGTAGAGCTGGT 1857
Qy 241 TTTTATTATAATATTATTTCAAGTTTGCAGATATTAGTATATAATTTGTTAAAGAAATGAA 300
Db 1858 TCTTTTATTATAATATTATTTCAAGTTTACAGATGTAATTTTGAATTTACAGAGATTGAA 1917
Qy 301 CAACACCAAGTGTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTT 360
Db 1918 CAACATCAAAATGTTCTTACTAGTGGGAAATCAGAGATTACTTTTAAAGGTAAGGATGTC 1977
Qy 361 GACCAATATCCTGCTACAGAAAGTATCAACAGAAATCTTTGATTTTAAACAAATA 420
Db 1978 GATCAATACCTCTGCTCAGGAAATGACACAGATCTCCTAATTAACATTAGAACTTAA 2037
Qy 421 TTTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTGCT 480
Db 2038 CTGTTAAATCAATTTATTAATGAACTGCTTTTGTCTAGCCACAGAAAGCGTCCA 2097
Qy 481 ATTTTACAGAGTTCATATTGTTATTAGTAAATCAATAAGATTTTAAAGCAGTAGCGACT 540
Db 2098 ATCTTACAGGTGTCATTTGCTTATCAGTCAAAATAAATCTTTTAAAGGCTGTTGCGACA 2157
Qy 541 GACTCTCATCTGATGAGCAAGGTTTAACTACTTTGGACAATCTTTCAGCAGATTGATG 600
Db 2158 GATTACACCGTATGAGTCAACGCACTTTCCAAATTAGAGAAATCGGCTAAATTTTATGAT 2217
Qy 601 GTAGTTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTAG 660
Db 2218 TTGGTTGTTCCAAGTAAATCTCTTCGAGAAATTTTTCGGCTGTTTTTACAGATGATTTGAA 2277
Qy 661 ACCGTTGAGGTATTTTCTTCCAAAGCCAAATCTTGTTCAGAAAGTGAACATTTCTTTT 720

Db 2278 ACTTAGAGGTTTCTTCTCAGATAGTCAAAATGTTATTTAGAGTGAAAAATATCAGCTTC 2337
Qy 721 TATACAGCCCTCTTTAGAGGAAAAATATCCCGATACAGACCGGTTTATTAATGACGAAATTT 780
Db 2338 TATACAGGTTTGTCTTGAAGGAAACTACCTGATACTGATCGCTCCTTAACATCAGTTT 2397
Qy 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCCGACCGTATGGAACGTCCTTCTTTG 840
Db 2398 GAAACCGAAATATCTTTAATACAAATGCTTTACGCCATGCTATGGAACGTCCTTATTTA 2457
Qy 841 ATTTCTAATGCTACTCAAAATGTTAGTAAAGCTTGAGATTACTTCAAAATCATATTTCA 900
Db 2458 ATTTGCAATGCACTCAGACGGTACTGTTGTTTGAATAATCAAAATGAGACAGCTTCA 2517
Qy 901 GCTCATGTTAACTCACCTGAGGTTTGAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 960
Db 2518 GCTCATGTTAACTCTCCAGGTTTGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2577
Qy 961 TCTGGTAGTGAATTAATCACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 1020
Db 2578 AAAGGTGATAGTTTAAATATTTAGCTTTAATCCAACTTACCTTAAATTTGAATCTTTTAAAGCA 2637
Qy 1021 ATTTAAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTCGACCAATTCACCCCTACA 1080
Db 2638 GTAAAAAGCGAAACAGTTACGATTCGATTTATTTCTCCAGTACGTCCTTACTTTTGACA 2697
Qy 1081 CCAGCGATGAGGAAGAAAGTTTATCCAAATTAATTAACAGGTAAGGTAAGGTAAGGTAAG 1134
Db 2698 CCTGGTGAAGATCTGAAGATTTCATCCAAATTAATTAACGTCCTGTCGTACTTAAC 2751

RESULT 14
AX602184
LOCUS AX602184 26144 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 113 from Patent WO02092818.
ACCESSION AX602184
VERSION AX602184.1 GI:28402053
KEYWORDS
SOURCE Streptococcus agalactiae
ORGANISM Streptococcus agalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1
AUTHORS Glaser, P., Rusniok, C., Chevalier, F., Frangeul, L., Lalioui, L.,
Zouine, M., Couve, E., Buchrieser, C., Poyart, C., Trieu-Cuot, P. and
Kunst, F.
TITLE Streptococcus agalactiae genome sequence, use for developing
vaccines, diagnostic tools, and for identifying therapeutic targets
JOURNAL Patent: WO 02092818-A 113 21-NOV-2002;
INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (FR)
FEATURES
source 1-26144
Location/Qualifiers
/organism="Streptococcus agalactiae"
/mol_type="unassigned DNA"
/db_xref="taxon:1311"

Query Match 56.5%; Score 641.2; DB 6; Length 26144;
Best Local Similarity 72.8%; Pred. No. 2.7e-108;
Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
Qy 1 ATGATTCATATTTCAATTAATCGCATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 10096 ATGATTCATATTTCAATTAATTAATAATTTCTTCTGATGCTCTAACCGTAAACCAACGA 10155
Qy 61 GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAAGTCACTTCT 120
Db 10156 GCTATTAGCATAAAATGCCATTCCTATTCTTTCATCAATAAAAAATGAAAGTCACTTCT 10215
Qy 121 ACAGAGTAACTTTAAACAGGCTCTAACCGTCAATATCAATTTGAAACACTATTCTCTGTA 180


```
Qy 661 ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATCTCTTTT 720
Db 33757 ACTGTAGAGGTTTTCTTCTTCAGATAGTCAAAATGTTATTTAGAGTGAATAATCAGCTTC 33816
Qy 721 TATACACGCTCTTAGAGAGGAAATTTATCCCATATACAGACCGTATTATTAATGACAGAAATTT 780
Db 33817 TATACACGTTTGCTTGAAGGAAACTACCTCATACTGATCGCTCTTAACCTAATCAGTTT 33876
Qy 781 GAGACGGAGGTGTTTTCATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTG 840
Db 33877 GAAACCGGAAATTTATCTTTAATACAAATGCTTTAGCCCATGCTATGGAACGCTGTTATTTA 33936
Qy 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAATCATATTTCA 900
Db 33937 ATTTGGAATGCAACTCAGAACGGTACTGTTGTTTAGAAAATTCAAAATGAGACAGTCTCA 33996
Qy 901 GCTCATGTAACTCACTCAGCTGGTAAAGCTAAACGAGGATTTAGATATTGTTAGTCAG 960
Db 33997 GCTCATGTAACTCTCCAGAAAGTTGGTAAAGTTAATGAGGAATGGATCTGTTAGTCTT 34056
Qy 961 TCTGGTAGTATTAACTATATCAGCTTCAATCCAACTTACCTTATGAGTCTTTAAAGCT 1020
Db 34057 AAAGGTGATAGTTTAAATATTAGCTTTAATCCAACTTACCTAATGAAATCTTTAAAGCA 34116
Qy 1021 ATTTAAAGTGAACAGTAAATTCATTTCTTATCACCAGTTCGACCATTCACCCATAACA 1080
Db 34117 GTAAAACGGAACAGATGTTACGATTTATTTCTCCAGTACGTCCTATTACTTTGACA 34176
Qy 1081 CCAGCGATGAGGAAGAAAGTTTATCAATTAATTTACACCACTGACGAAACAAAC 1134
Db 34177 CTGGTGAAGATCTGAAGATTTCAATCCAAATTAATTAATCTGTTGCTACTAAC 34230

RESULT 16
LOCUS SAG766843 143050 bp DNA linear BCT 05-NOV-2002
DEFINITION Streptococcus agalactiae NEM316 complete genome, segment 1.
ACCESSION AL766843 AL766843
VERSION AL766843.1 GI:23094426
KEYWORDS
SOURCE Streptococcus agalactiae NEM316
ORGANISM Streptococcus agalactiae NEM316
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
1 Glaser,P., Rusniok,C., Chevallier,P., Buchrieser,C., Frangeul,L.,
Zouine,M., Couve,E., Lalioui,L., Msadek,T., Poyart,C.,
Trieu-Cuot,P. and Kunst,P.
Genome sequence of Streptococcus agalactiae, a pathogen causing
invasive neonatal disease
Mol. Microbiol. 45 (6), 1499-1513 (2002)
22242508
PUBMED 12354221
REFERENCE
2 Glaser,P., Rusniok,C. and Frangeul,L.
Direct Submission
Submitted (31-MAY-2002) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE. E-mail: pglaser@pasteur.fr Phone: +33 (0)1 45 68
89 96, Fax: +33 (0)1 45 68 87 86
FEATURES
Location/Qualifiers
1..143050
/organism="Streptococcus agalactiae NEM316"
/mol_type="genomic DNA"
/strain="NEM316"
/serovar="serovar III"
/db_xref="taxon:211110"
165..1537
/gene="dnaA"
165..170
/gene="dnaA"
176..1537
/gene="dnaA"

gene
RBS
CDS
```

4188. .4391
 /gene="gbs00005"
 4188. .4391
 /gene="gbs00005"
 /notes="Hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD45650.1"
 /db_xref="GI:23094431"
 /db_xref="TrEMBL:Q8E720"
 /translation="MDNLRLAEIPDVSLEHVLAKTNEVKVERIYENKPLDLRKYNK
 IYWFIPRNIRISLLIILAILFL"
 4537. .5666
 /gene="gbs00006"
 4537. .5666
 /gene="gbs00006"
 4551. .5666
 /gene="gbs00006"
 /note="similar to hypothetical GTP binding protein"
 /codon_start=1
 /transl_table=11
 /product="Unknown"
 /protein_id="CAD45651.1"
 /db_xref="GI:23094432"
 /db_xref="GOA:Q8E7Y9"
 /db_xref="TrEMBL:Q8E7Y9"
 /translation="WALTAGIVGLPNVCKSLFNATKAGAEAAAYFPATIDPNVGVY
 EVDPERLOKLELTPPKTIVPTFEFDIAGIVKSGKEGLGNKFLANIREVDAIVH
 VRAFDDENWREQREDAFVDPIADIDITINLEILADLESINKRYARVERKARTQKD
 KESVAEFNVLOKIPVLEDGKSARTIETFEERKVVKGFLFTTKPVLYVANVEDQKD
 ADPDIDYNGIRAFAPETENAEEVWVVISARAEETSELDDDEKLEFLAIGLTSRGVDK
 LTRAYHLLGLTGYFTAGEKEVRAWTFKRGIKAPQAAIISHSDFERGFIRAVTMSYDD
 LIQYSEKAEVAGRLREEGKEYIVQDQIMEFENVV"
 join(5739. .5744, 5750. .6325)
 /gene="gbs00007"
 5739. .5744
 /gene="gbs00007"
 5750. .6325
 /gene="gbs0007"
 /note="similar to peptidyl-tRNA hydrolase"
 /codon_start=1
 /transl_table=11
 /product="Unknown"
 /protein_id="CAD45652.1"
 /db_xref="GI:23094433"
 /db_xref="GOA:Q8E7Y8"
 /db_xref="TrEMBL:Q8E7Y8"
 /translation="WVWMIVLGNPGSKYNDTKHNTGMAVDRIKDLVDNFTEDKRF
 KAEIGSPFINGEKTYFKPTFMNNSGIAVKALLTYNISIKDMLIIYDDLDMEVGKI
 RFQKSGAGHNGKSIHAIHLGTQEFDRIKVGIQRPNGRMTVINHVLGFPDRKNDIMI
 LNTLDKVDNAVYYLQNDNFQTKMYKNGLK"
 join(6322. .9819, 9893. .9909)
 /gene="gbs00008"
 6322. .9819
 /gene="gbs0008"
 /note="similar to transcription-repair coupling factor"
 /codon_start=1
 /transl_table=11
 /product="Unknown"
 /protein_id="CAD45653.1"
 /db_xref="GI:23094434"
 /db_xref="GOA:Q8E7Y7"
 /db_xref="TrEMBL:Q8E7Y7"
 /translation="MNIIFSQNKVVRWTHSGLVNRSQLVMGFSGAKATAIASAY
 EKLKIMDKVNTQTDSDLSIDIGEDNVYQFADDDVPAABFIFSSLDKSI
 SAURFLKDPKNGVLITISGLRLLPNEVFSKQKEIQAECVLDKLCNVLNGL
 YQKVSQVSPFGESQDILIDPEMTQEPYRLEFDEIDGROFDIDTOKSLQLE
 SVQISPADDIILQDADFERRAKKLEGLVYASEVORTYLSVLTSTENHFKHSDIRP
 LSIIFYEKWGLIDYIPEGTPLFVDDPQKIVDRNAKLIDLEIASLLTEDLQOQKSHSLN
 YFSDPKQLRQYQAPATFSPFNHKGILNKFEDKLHFTQTGMQEFNFQFPLVDENRYA
 KKSQATVLLQVDSQKLNLLQENLKEIGLDLISDKNDIVQKESQLIIVHLSNGFYFA
 DEKIVLITREIYHRRKVRKRSINAEALKYNELSDGVYVHNVHGVGKFLGIE
 DEIQIHRDYLITQONADRLISIPVEQIELLTKYVSADGKEPKINTLNDGRFFKAKO

```

QY 781 GAGACGAGGTTGTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTG 840
Db 2472 GAAACCGGAAATATCTTTAAATACAAATGCTTTACGCCATGCTATGGAACGTCCTTATTTA 2531
QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTCAAAAATCATATTTCA 900
Db 2532 ATTTGCAATGCAACTCAGAACGGTACTGTTGCTTTAGAAATTCAAAATGAGACAGTCTCA 2591
QY 901 GCTCATGTTAACTCACTCAGGAGTGGTAAAGGTAACGAGGATTTAGATATTTAGTTAGTCAG 960
Db 2592 GCTCATGTTAACTCTCCAGAAAGTGGTAAAGTAAATGAGGAATTCGACTACTGTTAGTCTT 2651
QY 961 TCTGCTAGTGAATTAACCTATCAGCTTCAATCCAACTTACCTTATTTAGCTCTTTAAAGCT 1020
Db 2652 AAAGTGATAGTTTAAATATTAGCTTTTAAATCCAACTTACCTTAAATGAACTTTTAAAGCA 2711
QY 1021 ATTTAAAGTGAACAGTAAATTAATTTCTTATCACCAGTTTCGACCAATTCACCCCTAACA 1080
Db 2712 GTAAAGCGAAGACAGTTACGATTCGATTTATTTCTCCAGTACGTCCTTACTTTGACA 2771
QY 1081 CCAGCGATAGGAAGAAAGTTTATCCAAATTAATACCAAGTACGAAACAAAC 1134
Db 2772 CCTGCTGAAGATACCTGAAGATTTATCCAAATTAATTAATCTCTGTTGCTACTAAC 2825

RESULT 17
CQ655073 349980 bp DNA linear PAT 02-FEB-2004
LOCUS Sequence 12030 from Patent WO0234771.
DEFINITION CQ655073
ACCESSION CQ655073
VERSION CQ655073.1 GI:41687950
KEYWORDS Streptococcus agalactiae
SOURCE Streptococcus agalactiae
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

REFERENCE
1 Telford, J., Masignani, V., Margarit y Ros, I., Grandi, G., Fraser, C. and Tettelin, H.
TITLE Nucleic acids and proteins from streptococcus groups a & b
JOURNAL Patent: WO 0234771-A 12030 02-MAY-2002;
Chiron S.p.A. (IT): THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
source
1..349980
/organism="Streptococcus agalactiae"
/mol_type="unassigned DNA"
/db_xref="taxon:1311"
/note="seq 10967 to big: splitted in-seq 10.967, from 0.000.001 to 0.349.980-seq 12.025, from 0.300.001 to 0.649.980-seq 12.026, from 0.600.001 to 0.949.980-seq 12.027, from 0.900.001 to 1.249.980-seq 12.028, from 1.200.001 to 1.549.980-seq 12.029, from 1.500.001 to 1.849.980-seq 12.030, from 1.800.001 to 2.149.980-seq 12.031, from 2.100.001 to 2.160.266"

ORIGIN
Query Match 56.5%; Score 641.2; DB 6; Length 349980;
Best Local Similarity 72.8%; Pred. No. 1.6e-108;
Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

QY 1 ATGATTCATTTTCGAATTAATCGCATATTTATTTCATGCTTTTAAATACAACTAAACGT 60
Db 333097 ATGATTCATTTTTCGAATTAATTAATTAATTTCTTTCGCACTGCTAACCGGTAAACCAACGA 333156
QY 61 GCTATTAGCACTAAATAGCCATTCCTATTCTTTCAATCAATCAATTAATTAATTAATTAATTAAT 120
Db 333157 GCTATTAGCATATAAATAGCGATTTCCATCTTCACTGTTTAAATTAAGATGACTAGG 333216
QY 121 ACAGAGATCACTTTAAACAGGCTTAAACCGGTCAAAATATCAATTAATTAATTAATTAATTAAT 180
Db 333217 GATGCTATCATTTTAAACGGGTCAAAATGGAACAAATTTCAATTAATTAATTAATTAATTAAT 333276
QY 181 AGTAATGAATATGCTGGTTTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240

```

```

Db 333277 TCAAAATGAAATGCGAGTTTACTAGTAACGAATCAAGCTCTATTTTGTTAGAGCTGGT 333336
QY 241 TTTTATTAATATATTTTCAAGTTTGGCAGATATTAGTATATAAATGTTTAAAGAAATTTGAA 300
Db 333337 TTTCTTTAATAATATTTTCAAGTTTACCAGATGTAACTTTTAGAATTTTACAGAGATTGAA 333396
QY 301 CAACACCAAGTTGTTTAAACCAAGTGGTAAATCAGAGATTACTTTAAAGGAAAGATGTT 360
Db 333397 CAACATCAAAATTTGTTTCTTACTAGTGGAAATCAGAGATTACTTTTAAAGGTAAGGATGTC 333456
QY 361 GACCAAGTATCTCTGTACAAAGAAAGTATCAACAGAAAAATCTTTGATTTTTAAACAAAA 420
Db 333457 GATCAATACCTCTGTCTGCAGAAATGACAACAGATCTCCATTAACATTTAGAAACTAAA 333516
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTCAAGAAGAGTCTGCT 480
Db 333517 CTGTTAAAAATCAATTTATTAATGAAACTGCTTTTGTCTAGCCAAAGAAAGCGGTCCA 333576
QY 481 ATTTTAAACAGAGATTCTATATTGTAATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
Db 333577 ATCTTAAACAGGTGTTTCAATTTGTTTATCAGTCAAAATAAATACTTTTAAAGCTGTTGCGACA 333636
QY 541 GACTCTCATCTGATGAGCCAACTTTAATCACTTTGGACAATCTTCAAGCAGATTGTATG 600
Db 333637 GATTCAACACCTGATGAGTCAACGCACTTCCAAATTAGAGAAATCGGCTAATAATTTTGAT 333696
QY 601 GTAGTTCTTCCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTGAG 660
Db 333697 TTGGTTGTTCCAAGTAAATCTCTTCGAGAAATTTTTCGGCTGTTTTTACAGATGATATTGAA 333756
QY 661 ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTTGTCAGAGTGAACACATTTCTTTT 720
Db 333757 ACTGTAGAGGTTTCTTCTCAGATAGTCAAAATGTTTATTAGAAGTGAATAATATCAGCTTC 333816
QY 721 TATACAGCTCTTTAGAGGAAATTTATCCGATATCAGACCGTTTATTATATGACAGAAATTT 780
Db 333817 TATACAGCTTTGTTGAAGGAAACTACCTCGATCTGATCGCTCTTAATCAATCAGTTT 333876
QY 781 GAGACGAGGTTGTTTTCATACCCAAATCCCTTCGCCACGCTATGGAACGTCCTTCTTTG 840
Db 333877 GAAACCGAAATTTATCTTTTAAATACAAATGCTTTTACGCCATGCTATGGAACGTCCTTATTTA 333936
QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAATCATATTTCA 900
Db 333937 ATTTGCAATGCAACTCAGAACGCTACTGTTCTGTTTGAAGAAATTCAAAATGAGACAGTCTCA 333996
QY 901 GCTCATGTTAACTCACTCAGGTTGGTAAAGTAAAGTAAATGAGGAATTCGATACCTGTTAGTCTT 960
Db 333997 GCTCATGTTAACTCTCCGAAAGTTGGTAAAGTAAATGAGGAATTCGATACCTGTTAGTCTT 334056
QY 961 TCTGCTAGTGAATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCT 1020
Db 334057 AAAGTGATAGTTTAAATATTTAGCTTTAATCCAACTTACCTTAAATGAAATCTTTAAAGCA 334116
QY 1021 ATTTAAAGTGAACAGTAAATTAATTTCTTATCACCAGTTTCAGCACTTCAACCTTAACA 1080
Db 334117 GTAAAAACGGAACAGTTACGATTTGATTTATTTTCTCCAGTACGTCCTTACTTTTGACA 334176
QY 1081 CCAGCGATGAGGAAGAAAGTTTATCCAAATTAATTAACCAAGTACGAAACAAAC 1134
Db 334177 CCTGCTGAAGATACCTGAAGATTTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 334230

```

```

RESULT 18
AX954527 349980 bp DNA linear PAT 08-JAN-2004
LOCUS Sequence 1373 from Patent WO03093306.
DEFINITION AX954527
ACCESSION AX954527
VERSION AX954527.1 GI:40783900
KEYWORDS Streptococcus pyogenes
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes

```

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE
AUTHORS Telford, J., Massignani, V., margarit y Ros, I., Grandi, G., Fraser, C.
and Tettelin, H.

TITLE Nucleic acids and proteins from streptococcus groups a b
JOURNAL Patent: WO 03093306-A 1373 13-NOV-2003;

Chiron SRL (IT); THE INSTITUTE FOR GENOMIC RESEARCH (US)
Location/Qualifiers

1. 349980
/organism="Streptococcus pyogenes"
/mol_type="unassigned DNA"

/db_xref="taxon:1314"
/note="seq 1373 to long, original length:
2.160.266-replaced by-seq 1373: from 0.000.001 to

0.349.980-seq 1374: from 0.300.001 to 0.649.980-seq 1375:
from 0.600.001 to 0.949.980-seq 1376: from 0.900.001 to

1.249.980-seq 1377: from 1.200.001 to 1.549.980-seq 1378:
from 1.500.001 to 1.849.980-seq 1379: from 1.800.001 to

2.149.980-seq 1380: from 2.100.001 to 2.160.266"

ORIGIN

Query Match 56.5%; Score 641.2; DB 6; Length 349980;
Best Local Similarity 72.8%; Pred. No. 1.6e-108;
Matches 826; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

QY 1 ATGATTCATATTTCAATTAATCCACATTAATTTATTCATGCTTTAAATACAATAAAAGCT 60
Db 1618 ATGATTCATATTTCAATTAATTAATAAATTTCTTCTGATGCTTAACGGTAACCAACGA 1677

QY 61 GCTATTAGCACTAAATGCTATCTTCTTATTCATCAATTAATAATGAAGTCACTTCT 120
Db 1678 GCTATTAGTCATAAATGCGATTCCTCAATCTTCACTGTTTAAATAGAGTACTAGG 1737

QY 121 ACAGGAGTAACTTTAAACGGGTCTAACGGTCAATATCAATTTGAAAAACACTATTCCTGTA 180
Db 1738 GATGCTATCATTTTAAACGGGTCAATATGACAAATTCATTTGAATACTATTCCTGCT 1797

QY 181 AGTAATGAAATGCTGGTTGCTAATTAACCTCTCAGAGCTATTTTATTAAGAAGCTAGT 240
Db 1798 TCMAATGAAATGCGAGTTTACTAGTAACGAATCCAGGCTCTATTTGTTAGAGCTGGT 1857

QY 241 TTTTATTAATTAATTTCAAGTTGCGAGATTTAGTATTAATTAAGTAAGTAAATTTGAA 300
Db 1858 TTCTTTATTAATTAATTTCAAGTTTACAGATGTAATTTAGATTTTACAGAGATTTGAA 1917

QY 301 CAACACCAAGTTGTTTAAACGGTGGTAAATCAGAGATTTACCTTAAAGGAAAAAGATGTT 360
Db 1918 CAACATCAATTTGTTCTTACTAGTGGAAAAATCAGAGATTTACTTTGAAAGGTAAGGATGC 1977

QY 361 GACCAATATCCTCGTCTACAGAGATTTGCGAGATTTTGCAGCAGTTTACAGAAAGTCGCT 420
Db 1978 GATCAATACCTCGTCTGAGGAATGACACAGATCTCCATTAACATTAAGAACTAAA 2037

QY 421 TTATTGAAGTCTATTTATTTGCTGAAACAGCTTTTGCAGCAGTTTACAGAAAGTCGCT 480
Db 2038 CTGTTAAATCAATTAATTAATGAAATGCTTTTCTGCTAGCCCAACAGAAAGCCGTC 2097

QY 481 ATTTTAAACAGAGTTTCAATTTAGTATTAAGTAAATCATTAAGATTTTAAAGCAGTAGCACT 540
Db 2098 ATCTTAAACAGGTTTCAATTTGTTATTCAGTCAAAATTAATTAATTAAGGCTGTTGCGACA 2157

QY 541 GACTCTATCTGATGAGCCAAAGCTTTAATCACTTTGGACAATATCTTCCAGCAGATTTGATG 600
Db 2158 GATTACACCCGTATGATGATCAACGCACTTTCCAAATGAGAAATCGGCTAATATTTGAT 2217

QY 601 GTAGTTCTTCAAGTAAATCTTTGAGAGATTTTTCAGCAGTATTTACAGATGATTTGAG 660
Db 2218 TTGGTTGTTCCAGTAAATCCCTTCGAGAAATTTTCGGCTGTTTTTACAGATGATTTGAA 2277

QY 661 ACCGTTGAGGTTATTTTCTCACCAGCCAAATCTTGTTCAGAAAGTGAACACTTCTTTT 720
Db 2278 ACTGTAGAGGTTTCTTCTCAGATAGTCAATGTTTATTAGAGTGAATAATATCAGCTTC 2337

QY 721 TATACAGCCTCTTAGAAGGAATTTATCCGATACAGACCGTTTATTATTAATGACAGATTT 780
Db 2338 TATACAGCTTTGCTTTGAAGGAATTTATCCCTGATGATCGCTCTTAATCAATCAGTTT 2397

QY 781 GAGACGGAGTTGTTTTCATATCCCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTG 840
Db 2398 GAAACCGAAATTTATCTTTTAATACAATGCTTTACGCCATGCTATGGAACGTCCTTATTA 2457

QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
Db 2458 ATTTGCAATGCAAACTCAGAACGTTACTGTTGTTTAGAAATTCAAAATGACAGACTCTCA 2517

QY 901 GCTCATGTTAACTCACCTGAGTGGTAAAGTAAACGAGGATTTAGATTTAGTTAGTCAG 960
Db 2518 GCTCATGTTAACTCTCCAGAGTTGGTAAAGTTAATGAGGAATGGATACTGTTAGTCTT 2577

QY 961 TCTGTAGTGAATTTAACTATCAGCTTCAATCCAACTTACCTTATTTGAGTCTTTAAAAAGCT 1020
Db 2578 AAAGTGATAGTTTAAATATTTAGCTTTAATCCAACTTACCTTAAATGAACTTTAAAAAGCA 2637

QY 1021 ATTAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTGACCAATTCACCTTAACA 1080
Db 2638 GTAAAGCGAAACAGTTACGATTCGATTTATTTCTCCAGTACGTCATTTACTTTTGACA 2697

QY 1081 CCAGCGATGAGGAAGAAAGTTTATCCAAATTAATTAACAGTACGAAACAAAC 1134
Db 2698 CCTGGTGAAGATAGTGAAGATTTTATCCAAATTAATTAATCTCTGTTGCTACTAAC 2751

RESULT 19

AX571491

LOCUS

AX571491

DEFINITION

Sequence 4703 from Patent WO02077021.

AX571491

VERSION

AX571491.1

KEYWORDS

GI:26003814

SOURCE

Streptococcus pneumoniae

Streptococcus pneumoniae

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Streptococcus.

REFERENCE

1

Massignani, V., Tettelin, H. and Fraser, C.

Streptococcus pneumoniae proteins and nucleic acids

Patent: WO 02077021-A 4703 03-OCT-2002;

Chiron Spa (IT); THE INSTITUTE FOR GENOMIC RESEARCH (US)

Location/Qualifiers

1. .1134

/organism="Streptococcus pneumoniae"

/mol_type="unassigned DNA"

/db_xref="taxon:1313"

ORIGIN

Query Match

53.6%; Score 608.2; DB 6; Length 1134;

Best Local Similarity 71.1%; Pred. No. 5.9e-102;

Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 1 ATGATTCATATTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATTAACACTTAAACGT 60

Db 1 ATGATTCATATTTCAATTAATTAATAAATTTATTTCTCAAGCATTTAAATACTACTAAGAGA 60

QY 61 GCTATTAGCACTTAAATAATGCAATTCCTATTTCTTTCATCAATAAAATTTGAAAGTCACTTCT 120

Db 61 GCTATTAGTCTTAAATAATGCAATTCCTATTTTATCAACAGTAAATAATTTGACGTGACCAAT 120

QY 121 ACAGAGTAACTTTTAAACAGGCTTAAACGTCAAATATCAATTTGAAACACTATTTCTGTA 180

Db 121 GAGGATTAATTTTAAATTTGTTTCAATTTCAATTTCAATTTCAATTTTATTCTCAA 180

QY 181 AGTAATGAAATGCTGGTTTCTGCTAATTTACCTCTCAGAGCTATTTTATTAGAGGCTAGT 240

Db 181 AAAATGAGATGCTGGTTTCTGTTAATTTACTTTCTTTAGGTTGATCTCTTCTTGAAGCTTCT 240

QY 901 GCTAAGTAACTACCTGAGGTTGGTAAGTAAACGAGGATTAGATATTGTTAGTCAG 960
 Db |||||
 QY 901 GCCCATGTTCACTCTCCAGAAAGTTGGTAAAGTAAACGAGGAAATCGATACGAGTT 960
 Db |||||
 QY 961 TCTGTAGTAACTAACTATCAGCTTCAATCAACTTACCTTATGAGTCTTTAAAGCT 1020
 Db |||||
 QY 961 ACTGTGAGATTTGACCAATAGTTTCAACCAACTTACTTGTATTTCTTAAAGCT 1020
 Db |||||
 QY 1021 ATTAAGTGAACAGTAAATTCATTTCTATCACCAGTTGACCACTTACCCCTAACA 1080
 Db |||||
 QY 1021 TTAATAGCGAAAGGAGCAATAGCTTTATCTCAGCTGTTGTCATTTACTCTTGTG 1080
 Db |||||
 QY 1081 CCAGCGATGAGGAAGATTTTATCCATTAATACACAGTACGACAAA 1133
 Db |||||
 QY 1081 CCAGCAGATATGACGAAGACTTCATGACGCTCATACACAGTTCGTACAAA 1133
 Db |||||

RESULT 21
 AE007318 14418 bp DNA linear BCT 31-AUG-2001
 LOCUS Streptococcus pneumoniae TIGR4 section 1 of 194 of the complete genome.
 ACCESSION AE007318 AE005672
 VERSION AE007318.1 GI:14971464
 KEYWORDS
 SOURCE Streptococcus pneumoniae TIGR4
 ORGANISM Streptococcus pneumoniae TIGR4
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 1 (bases 1 to 14418)
 Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickinson, T., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
 Complete genome sequence of a virulent isolate of Streptococcus pneumoniae
 Science 293 (5529), 498-506 (2001)
 21357209
 11463916
 2 (bases 1 to 14418)
 Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Gesuwan, P., Hickey, E.K., Holt, I.E., Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
 Direct Submission
 Submitted (29-Jun-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers
 1. 14418
 /organism="Streptococcus pneumoniae TIGR4"
 /mol_type="genomic DNA"
 /strain="TIGR4"
 /db_xref="taxon:170187"
 197. 1558
 /gene="SP0001"
 197. 1558
 /gene="SP0001"
 /note="identified by match to TIGR protein family HMM TIGR00362"
 /codon_start=1
 /transl_table=11
 /product="chromosomal replication initiator protein DnaA"
 /protein_id="AAK74194.1"

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Gesuwan, P., Hickey, E.K., Holt, I.E., Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
 Direct Submission
 Submitted (29-Jun-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers
 1. 14418
 /organism="Streptococcus pneumoniae TIGR4"
 /mol_type="genomic DNA"
 /strain="TIGR4"
 /db_xref="taxon:170187"
 197. 1558
 /gene="SP0001"
 197. 1558
 /gene="SP0001"
 /note="identified by match to TIGR protein family HMM TIGR00362"
 /codon_start=1
 /transl_table=11
 /product="chromosomal replication initiator protein DnaA"
 /protein_id="AAK74194.1"

/db_xref="GI:14971465"
 /translation="MKKQFNRILFEAQLRTRMYDFAIOAELIKVENVATIFL
 PREMEMEXQLKDIIVVAGFEIYDABITPHYITPKQDTSQVEATNLTVNYS
 PKLVIPYSDTKLEKTYFDNFIOGDNVWVAALAVSEDLATYNPLFYGGPGLG
 KTHLNAGNLIKNI PNARVKYIPAESFINDELDHLRLGEMKEKFTYRSIDLLID
 DIQSLGKVKVQTOBEFFNALHDKQKQIVLTSDSPKHLGLEERLVTRFSGWLTQ
 TITPDEFTRIALQSKTEHGLGYNFSQSDTSLAGQFSDNVRLDEGAINDTLLIARVK
 KIKITIDIAABARAKQDVSOMLVPIDKIQTEVGNFYGVSIKEMKSSRLQNI
 ARQVAYLURELTDNSLPKIGKEFGGKHDTTIIHAHAKIKSLIDQDDNLRLEISIKK
 KIK"
 1717. 2853
 /gene="SP0002"
 1717. 2853
 /gene="SP0002"
 /note="identified by match to TIGR protein family HMM TIGR00663"
 /codon_start=1
 /transl_table=11
 /product="DNA polymerase III, beta subunit"
 /protein_id="AAK74195.1"
 /db_xref="GI:14971466"
 /translation="MIHFSINKLFLQALNTTKRAISSKNAIPILSTVKIDVTNEGIT
 LIGNGQISIEFNISQKNEADGLITSLGSLLEASFNNVSSLPDVTLDKFEION
 QVLTSGKSEITLKGSEQVPRIOETASSTPLILETKLLKIIINLEPAFASQESRP
 ILTGHFVLSQHKELTVAQDTSRLSKQKLTLEKNSDDVDVITPSRLRESAVFTDD
 IETVEIFPANNOLFRSENISFVTLLEGNYPDRLIPTDNTTITFNVNLRQSMH
 RARLSSATONGTVKLEIKGVVSAHVHSEVKVNEEDTDQVTEGDLTISFNPTYL
 IDSLKALNSEKVTISFISAVRPFTLVFADTDEDFMQLITPVTN"
 2864. 3112
 /gene="SP0003"
 2864. 3112
 /gene="SP0003"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAK74196.1"
 /db_xref="GI:14971467"
 /translation="MSLARLFYDILIEKESSMYOVGNFVEMKKSHACTIKSTGKCAN
 RWITRVGADIKIKSCNCEHVVMGRYDFERRNKLID"
 3196. 4311
 /gene="SP0004"
 3196. 4311
 /gene="SP0004"
 /note="identified by match to PFAM protein family HMM PF01018"
 /codon_start=1
 /transl_table=11
 /product="GTP-binding protein"
 /protein_id="AAK74197.1"
 /db_xref="GI:14971468"
 /translation="WALTAGIVGLPNVGKSTLFLNATKAGAAANYPFATIDPNVGMV
 EYVDERLOKLTMTITPKTVPVTFEFTDIAGIVKGSKGEGKFLANIREVDAIVH
 VVAPDENVMSEQREDAFVDPDLADITINLELILADESVNKRVARVERKQATQKH
 KESVAFENVLQIKVLEKDSARTIEFTDESKVKVGLFLITLTPVIVANVDEVDK
 SPFSDIDYVKQIREFAATNAEVVISAPEEIEISLANDKKFELEAIGLTFESGVK
 LVRYAGHLGLGTYTAGKEVRAWTFKRGKAPQAAGIIHSDPEKGFTRAVTMSYED
 LVKYGSEKAVKAGRLREGEKEYIVQDGDIMEFRNV"
 4382. 4951
 /gene="SP0005"
 4382. 4951
 /gene="SP0005"
 /note="identified by match to PFAM protein family HMM PF01195"
 /codon_start=1
 /transl_table=11
 /product="peptidyl-tRNA hydrolase"
 /protein_id="AAK74198.1"
 /db_xref="GI:14971469"
 /translation="MTKLLVGLGNPGDKYFETKHNVGFMILDLQAKQNVTFTHDKIF
 QADLASFLNGEKIYLVKPTTFMSESGKAVHALLTYGLDIDLLIIYDDLDMEVGKI
 RLRAKSGAGGAGIKSIIOHIGTOVFNVRKIGIKGPKNGMSVHHVLSKFDREDDYIGI
 LQSDVKDDSVNYLQENKFEKTMQRYNG"

gene
 CDS
 gene
 CDS
 gene
 CDS
 gene
 CDS

```
gene
CDS
4952..8461
/gene="SP0006"
4952..8461
/gene="SP0006"
/notes="identified by match to TIGR protein family HMM
TIGR00580"
/codon_start=1
/transl_table=11
/product="transcription-repair coupling factor"
/protein_id="AAK74199.1"
/db_xref="GI:14971470"
/translation="MVTLLDLPSENDQIKKHONLTDKKROLILGLSTSKALAIASS
LEKEDRIVLLSTYGEAEGLVSDLLSILGEELVPELVDDAPMVEFLMSSOBKILSRV
EALRPLTSGGILVNCIAASRLILSPNAPKDSIVKISVGEYDQHAFIHLKENG
YKRVQVQTFGEFSURGDILDIPEISQLEPCRIEFDIIGIRSFVETQLSKENKT
ELTIPSPADMLTREDYQKQSALEKQISKTLSPILSKYLEIILSSPHOKSHADSRK
FLSLCYDTMTLVFDIYEDTPIFFDYQKLNQYEVFERDLAQFTLEQLONKSPSDM
QVFSDEIOLYKQSPVTFPSNLQKGLGKPKDKIYQFNQYPMQBFNQSFLKEIER
YKMDYTIILQSSNSMGSKTLEDMLEEYQIKLDSRDKNI CKESVNLJEGNLRGPHF
VDEKILLITEHIFOKKLRPRROHVNAERLNDYNLEKGDYVHHIHIGIYQLGI
ETIEIKGHRDYSVQYQNGQISIPVQIHLLSKYISSDGKAPKLNKLDGHRFKAK
KRVNQVEDIADDLIKLYSERSQLGFASADDDQDADFDPYVETDDQLRSTEEI
KRDMAOSPMDRLVGDVGFQTEVAMRAAFKAVNDHKQVVLVPTVLAQOHTNPK
ERFONFAVNIQVLSRFSKQETATLEKLNKQVDDILGTHRVLSKDVVEADLGLMII
DEQRFVGHKETLKEALKQVDVLTATPIPRTLHMSMLGIRDLSVIETPTNRPV
QTYLEKNDSDVRDLVREMERGGQYIILYINKVDITIVQVSELQELIPEASIGYVHR
MSEVQLENTLDFIEGQYDILVTVTIITETGVDIENANTLFITENADHMLGSLYQLRGR
VGRSNRIAYVLMYRPEKSI SEVASEKRLEAIKGFTELSGFKIAMRDLISIRGAGNLLG
KSGQFIDVSGFVYLSQLEAEAKRNGNANANRTKGNABLILQIDAYLPTVYSQD
RKIRIKYKIQIDNRNVVEELOEELIDRFGEYDVAVVALEIGLVKSYLDKVFQVRY
ERKNKTIQIEKVYQRLFLAQDYFKALSVTNLKGAIENKGLMELVPDQVKNQDYEI
LEGLIFGESLLEIKESKEENSI"
8519..8785
/gene="SP0007"
8519..8785
/gene="SP0007"
/codon_start=1
/transl_table=11
/product="84 domain protein"
/protein_id="AAK74200.1"
/db_xref="GI:14971471"
/translation="NRLEKYLKVSRIIKRRTVAKEVADKGRIVKNGILAKSSTDLLKNV
DQVEIRFGNKLVLVLEKDKSTKEDAGAGHYEIISETRVEENV"
8778..9146
/gene="SP0008"
8778..9146
/gene="SP0008"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK74201.1"
/db_xref="GI:14971472"
/translation="MSKNIYVNLNSFIQNEYORREYLMKERQKRFRMGVLLIMLL
FILPFLNAQSVQQLLQRRQLADLQTYQLISDEKQETAFATKLKDEDAAYKYTRA
KIYYSKREKVYTIPLLQR"
9151..9273
/gene="SP0009"
9151..9273
/gene="SP0009"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK74202.1"
/db_xref="GI:14971473"
/translation="MENLDDVIEQLFSLSDLEKLEBLADKNQLLRLOBEKEKRNA"
9266..10534
/gene="SP0010"
9266..10534
/gene="SP0010"
/notes="similar to PID:683698; identified by sequence
similarity; putative"
```

```

/codon_start=1
Query Match 53.6%; Score 608.2; DB 1; Length 14418;
Best Local Similarity 71.1%; Pred. No. 3.5e-102;
Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;
Qy 1 ATGATTCAATTTTCAATTAATCGCACATTATTATTTCATGCTTTTAAATACAACTAAACGT 60
Db 1717 ATGATTCAATTTTCAATTAATTAATAAATTTATTCTTACAAAGCAATTAATACTACTAAGAGA 1776
Qy 61 GCTATTAGCAGTCAAAAAATGCCATTCTCTATTCTTTTCATCAATAAAAAATTTGAAGTCACTTCT 120
Db 1777 GCTATTAGTCTTAAAAATGCCATTCTCTATTATTATCAACAGTAAAAATTTGACGTGACCAAT 1836
Qy 121 ACAGAGTAACCTTTAAACAGGCTTAACCGTCAAAATATCAATTTGAAAAACATTTCTCTGTA 180
Db 1837 GAAGGTATTACTTTAAATGGTTCAAAATGGTCAAAATTTCAATTTGAAAAATTTTATTCTCAA 1896
Qy 181 AGTAATGAATGCTGGTTTCTTAATTAACCTCTCCAGGAGCTATTATTATTAGAAGCTAGT 240
Db 1897 AAAAATGAAGATGCTGGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1956
Qy 241 TTTTATTATAATATTATTCAAGTTTGCAGATATTAGTATAAATGTTTAAAGAAAAATTTGAA 300
Db 1957 TTTTATTCAATAGTATCTAGTTTACCTGATGTAATCTCTGATTTTAAAGAAAAATTTGAA 2016
Qy 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAAGATGTT 360
Db 2017 CAAAATCAATTTGTTTAAACAGTGGCAATCAGAAAAATACCTTTAAAGGAAAAAGATAGC 2076
Qy 361 GACCATGATCTCGTCTACAGAGATATCAACAGAAAAATCTTTGATTTTAAAGAAAAA 420
Db 2077 GAACAATATCACGAATCCAGAAAAATTTTCAGCAAGCACTCTCTTTAATACTTTGAAAAA 2136
Qy 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTTCAGGAGCTTTTACAAAGAAAGTCGTCT 480
Db 2137 TTACTCAAGAAAAATTAATTAAGAAACAGCTTTTTCGTCAGGTACACAGAGAGTGTCTCG 2196
Qy 481 ATTTTAAACAGGATTCATATTGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCA 540
Db 2197 ATTTTAAACAGGTGCTCACTTCGTATTGATCAACACAAAGAGTGTAAAAACAGTTGCAACA 2256
Qy 541 GACTCTCATGATGAGCCCAAGTTTATCATCTTTGGCAATATCTTTCAGCAGATTTTGATG 600
Db 2257 GACTCTCATGCTTAAGCCAGAAAAAATTTGACTCTTTGAAAAAATAGTGATGATTTGAT 2316
Qy 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGCAGATTTTACAGATGATATTGAG 660
Db 2317 GTCTGATTTCTAGCGGTCTCTACGCGAAATTTTCAGCGGTATTTACAGATGATATCGAA 2376
Qy 661 ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
Db 2377 ACTGTAGAGATTTTCTTTTGGCCAAATAACCAAAATCTCTTTTAGAAGCGAAAAATATTAGCT 2436
Qy 721 TATACAGCCCTCTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTATGACAGAAATTT 780
Db 2437 TATACTGCTCTCTAGAAGGAAACTATCTCTGATACAGATCGCTTGATTTCCACAGACTTT 2496
Qy 781 GAGACGGAGGTGTTGTTTCAATACCAATCCCTTCGCCACGCTATTCGAAACGCTCTCTTTG 840
Db 2497 AACACTATATTACTTTTAAATGTTGGTAACTTACGCCAGTCAATGAGAGGTGCGCTCTTT 2556
Qy 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
Db 2557 TTATCAAGTGCAGCTCAAAATGGTACTGTGAAACTTTGAAATTAAGGATGGGGTGTGTAGC 2616
Qy 901 GCTCATGTTAATCACTACCTGAGGTGGTAAAGTAAACAGGAGTTTATGATATTGTTAGTCAG 960
Db 2617 GCCCATGTTCTCTCCAGAAAGTTGGTAAAGTAAACGAAAGAAATCGATATGATCAGGTT 2676
Qy 961 TCTGTAGTGAATTTAACTATCATGCTTCAATCCAACTTACCTTATTGAGTCTTTTAAAGCT 1020
Db 2677 ACTGTGGAAGATTTGACCAATTAGTTTCAACCCAACTTACTGATTTGATTTCTTAAAGCT 2736
```


QY 1021 ATTAAAGTGAACAGTAAATAATTCATTCTTATCACCAGTTTCGACCATTCACCCCTAACA 1080
 DB 2737 TTAATATAGCAAAAGGTGACTATTAGCTTTATCTCAGCTGTCTGCTCAATTACTCTTG 2796

QY 1081 CCAGCGATGAGGAAGAAGTTTATCCAAATTAATACACAGTACGAACAAA 1133
 DB 2797 CCAGCAGATGACGAAGACTTCATGAGCTCATTACACAGTTCGTACAAA 2849

RESULT 22
 CQ788929 21338 bp DNA linear PAT 29-MAR-2004
 DEFINITION Sequence 20 from Patent EP1400592.
 ACCESSION CQ788929
 VERSION CQ788929.1 GI:45822497
 KEYWORDS
 SOURCE unidentified
 ORGANISM unclassified.

REFERENCE 1
 AUTHORS Kunsch, C.A., Choi, G.H., Dillon, P.J., Rosen, C.A., Barash, S.C.,
 Fannon, M. and Dougherty, B.A.
 TITLE Streptococcus pneumoniae polynucleotides and sequences
 JOURNAL Patent: EP 1400592-A 20 24-MAR-2004;
 HUMAN GENOME SCIENCES, INC. (US)
 FEATURES Location/Qualifiers
 source 1. 21338
 /organism="unclassified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32644"

ORIGIN

Query Match 53.6%; Score 608.2; DB 6; Length 21338;
 Best Local Similarity 71.1%; Pred. No. 3.2e-102;
 Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCAGATTTATTCATGCTTTAAATACAACTAAACGT 60
 DB 7076 ATGATTCATTTTCAATTAATTAATAATTTATTTCTACAAGCAATTAATCTACTAAGAGA 7135

QY 61 GCTATTAGCACTAAATGCCATTCCTATTCTTCAATCAATAAATAATGAAGTCACTTCT 120
 DB 7136 GCTATTAGTCTTAAATGCCATTCCTATTCTTATCAACAGTAATAATTCAGTGCACCAAT 7195

QY 121 ACAGGAGTAACCTTACAGGCTCTACAGCTCAATATCAATTTGAAGAACACTATTCCTGTA 180
 DB 7196 GAAGGATTAATCTTAAATGCTTCAATGGTCAATTTCAATTTGAAGAAATTTTATTTCTCAA 7255

QY 181 AGTAATGAAATGCTGGTTGCTTAATTAATCTCTCCAGGAGCTATTATTTATGAAGCTAGT 240
 DB 7256 AAAAATGAAGATGCTGGTTGTTAATTAATCTTCTTTAGTTTCATCTTCTTGAAGCTTCT 7315

QY 241 TTTTATTATTAATTTATTTCAAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATTCGAA 300
 DB 7316 TTCTTTATCAATGTAGTACTAGTTTACGTATGATGTAATCTTGAATTTTAAAGAAATTCGAA 7375

QY 301 CAACACCAAGTTGTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAGATGTT 360
 DB 7376 CAATATCAATGTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAGATGAGC 7435

QY 361 GACAGTATCTCTGCTCAAGAAGTATCAACAGAAATCTTTGATTTTAAACACAAA 420
 DB 7436 GAACAATATCCCAAGTAATCAAGAAATTTCAAGAGCACTCTTTAATATCTTGAACAAAA 7495

QY 421 TTTTGAAGTCTATTATTGCTGAACAGCTTTTCAGCCAGCTTTTCAAGAAAGTTCGCTCT 480
 DB 7496 TTACTCAGAAATTTATTAATGAACAGCTTTTCTGCAAGTACACAGAGAGTTCGCTCG 7555

QY 481 ATTTTACAGGAGTTCAATTTGATTAATGAATATCAATAAGATTTTAAAGCAGTATGAGCACT 540
 DB 7556 ATTTTACAGGAGTTCACCTTCTGATTTAGTCAACACAGAGTTTAAACAGTTCGCAAC 7615

QY 541 GACTCTCATCGTATGAGCAACGTTTAAATCACTTTTGGACAATFACCTTCAGCAGATTTGATG 600
 DB 7616 GACTCTCATCGCTAAGCGAGAAAAAATGACTCTTTGAAAAAATAGTATGATGATTTGAT 7675

QY 601 GTAGTTCTTCCAAAGTAAATCTTTGAGAGATTTTTCAGCAGTATTTTACAGATGATATGAG 660
 DB 7676 GTGCTAATTTCCAGCGGTTCTCTACGCGAATTTTTCAGCGGATTTTACAGATGATATCGAA 7735

QY 661 ACCGTTGAGGTATTTTCTCCACCAAGCCAAATCTGTTTCAAGAGTGAACACATTTCTTTT 720
 DB 7736 ACTGTAGAGATTTTCTTTGCGCAATTAACCAATCTCTTTTGAAGCGAAAAATATTAGCTTC 7795

QY 721 TATACACGCTCTTGAAGGAAATTTATCCCGATACAGACCGTTTATTAATACAGAAATTT 780
 DB 7796 TATACTCGTCTCTAGAAGGAACTATCTCTGATACAGATCGCTTGTGATTTCCACAGACTTT 7855

QY 781 GAGACGGAGGTTGTTTTCATACCAATCCCTCGCCACGCTATGGAACGCTCTTCTTG 840
 DB 7856 AACACTACTATTACTTTTAAATGTGTAAACTTACGCCAGTCAATGGAGCGTGGCCGCTT 7915

QY 841 ATTCTTAATGCTACTCAAAATGGTACTGTTAAAGCTTGAGATTACTCAAAATCATATTTCA 900
 DB 7916 TTATCAAGTGCAGCTCAAAATGGTACTGTTGAACTTGAATTAAGGATGGGTTGTTAGC 7975

QY 901 GCTCATGTTAACTCACCTGAGGTTGGTAAGGTAACGAGGATTTAGATATTTGTTAGTCAG 960
 DB 7976 GCCCATGTTCACTCTCCAGAGTTGGTAAAGTAACGAGAAATCGATCTGATCAGGTT 8035

QY 961 TCTGGTAGTATTTAACTATACAGCTTCAATCCAACTTACCTTATGAGTCTTTTAAAGCT 1020
 DB 8036 ACTGGTGAAGATTTGACCATTTAGTTTCAACCCCACTTACTTGAATGATCTCTTAAAGCT 8095

QY 1021 ATTAAAGTGAACAGTAAATAATTCATTCTTATCACCAGTTCGACCATTCACCCCTAACA 1080
 DB 8096 TTAATAGCAAAAGGTGACTATTAGCTTTATCTCAGCTGTTCTGCTCAATTTACTCTTG 8155

QY 1081 CCAGCGATGAGGAAGAAGTTTATCCAAATTAATACACAGTCAATACCCAGTTCTGTACAA 1133
 DB 8156 CCAGCAGATCTGACGAAGACTTCAATGACGCTCAATACCCAGTTCTGTACAA 8208

RESULT 23
 AR218788 21338 bp DNA linear PAT 25-SEP-2002
 LOCUS Sequence 20 from patent US 6420135.
 DEFINITION AR218788
 ACCESSION AR218788
 VERSION AR218788.1 GI:23319722
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 21338)
 AUTHORS Kunsch, C.A., Choi, G.H., Dillon, P.S., Rosen, C.A., Barash, S.C.,
 Fannon, M.R. and Dougherty, B.A.
 TITLE Streptococcus pneumoniae polynucleotides and sequences
 JOURNAL Patent: US 6420135-A 20 16-JUL-2002;
 FEATURES Location/Qualifiers
 source 1. 21338
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 53.6%; Score 608.2; DB 6; Length 21338;
 Best Local Similarity 71.1%; Pred. No. 3.2e-102;
 Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCAATTAATCGCAGATTTATTCATGCTTTAAATACAACTAAACGT 60
 DB 7076 ATGATTCATTTTCAATTAATTAATAATTTATTTCTACAAGCAATTAATCTACTAAGAGA 7135

QY 61 GCTATTAGCACTAAATGCCATTCCTATTCTTCAATCAATAAATAATGAAGTCACTTCT 120
 DB 7136 GCTATTAGTCTTAAATGCCATTCCTATTCTTATCAACAGTAATAATTCAGTGCACCAAT 7195

QY 121 ACAGGAGTAACCTTACAGGCTCTACAGCTCAATATCAATTTGAAGAACACTATTCCTGTA 180
 DB 7196 GAAGGATTAATCTTAAATGCTTCAATGGTCAATTTCAATTTGAAGAAATTTTATTTCTCAA 7255

QY 181 AGTAATGAAATGCTGGTTGCTTAATTAATCTCTCCAGGAGCTATTATTTATGAAGCTAGT 240
 DB 7256 AAAAATGAAGATGCTGGTTGTTAATTAATCTTCTTTAGTTTCATCTTCTTGAAGCTTCT 7315

QY 241 TTTTATTATTAATTTATTTCAAGTTTGCAGATATTAGTATATAATGTTTAAAGAAATTCGAA 300
 DB 7316 TTCTTTATCAATGTAGTACTAGTTTACGTATGATGTAATCTTGAATTTTAAAGAAATTCGAA 7375

QY 301 CAACACCAAGTTGTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAGATGTT 360
 DB 7376 CAATATCAATGTTTAAACAGTGTAAATCAGAGATTTACCTTAAAGGAAAGATGAGC 7435

QY 361 GACAGTATCTCTGCTCAAGAAGTATCAACAGAAATCTTTGATTTTAAACACAAA 420
 DB 7436 GAACAATATCCCAAGTAATCAAGAAATTTCAAGAGCACTCTTTAATATCTTGAACAAAA 7495

QY 421 TTTTGAAGTCTATTATTGCTGAACAGCTTTTCAGCCAGCTTTTCAAGAAAGTTCGCTCT 480
 DB 7496 TTACTCAGAAATTTATTAATGAACAGCTTTTCTGCAAGTACACAGAGAGTTCGCTCG 7555

QY 481 ATTTTACAGGAGTTCAATTTGATTAATGAATATCAATAAGATTTTAAAGCAGTATGAGCACT 540
 DB 7556 ATTTTACAGGAGTTCACCTTCTGATTTAGTCAACACAGAGTTTAAACAGTTCGCAAC 7615

QY 121 ACAGAGTAACTTTAACAGGCTTAACGGTCAAAATATCAATGAAACACATATTCCTGTA 180
 DB 7196 GAAGGTATTAATTAATGTTTCAAAATGGTCAAAATTTCAATTTGAAATTTTATTTCTCAA 7255
 QY 181 AGTAATGAAATGCTGGTTGCTAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 DB 7256 AAAAATGAAGATGCTGGTTGTTAATTAATCTTTTAGGTTCAGTCCCTCTTGAAGCTTCT 7315
 QY 241 TTTTATTAATTAATTAATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTGAA 300
 DB 7316 TTTTATTAATTAATTAATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTGAA 7375
 QY 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAAAGATGTT 360
 DB 7376 CAAAATCAAAATGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAAAGATGAGC 7435
 QY 361 GACCAAGTATCTCGTCTCAAGAAAGTATCAACAGAAATTCCTTTTCATTTTAAAGAAACAAA 420
 DB 7436 GAACAAATATCCAGAAATCAAGAAATTTTCAGCAAGCTCTCTTTAATTAATTTGAAACAAA 7495
 QY 421 TTTTGAAGTCTATTAATTTGCTGAAACAGCTTTTTCAGCAGGTTTCAAGAAAGTCTGCT 480
 DB 7496 TTTTGAAGTCTATTAATTTGCTGAAACAGCTTTTTCAGCAGGTTTCAAGAAAGTCTGCT 7555
 QY 481 ATTTTAAACAGAGTTTCATATTTGTTTAAAGTAAATCAATGAAATTTTAAAGCAGTACGACT 540
 DB 7556 ATTTTAAACAGAGTTTCATATTTGTTTAAAGTAAATCAATGAAATTTTAAAGCAGTACGACT 7615
 QY 541 GACTCTCATGCTATGAGCAAGCTTTTAAATCACTTTTGGCAATACATTTTCAGCAGATTTGATG 600
 DB 7616 GACTCTCATGCTATGAGCAAGCTTTTAAATCACTTTTGGCAATACATTTTCAGCAGATTTGATG 7675
 QY 601 GTAGTTCTTCCAGTAAATCTTTGAGAAATTTTCAGCAGTATTTTACAGATGATATTGAG 660
 DB 7676 GTAGTTCTTCCAGTAAATCTTTGAGAAATTTTCAGCAGTATTTTACAGATGATATTGAG 7735
 QY 661 ACCGTTGAGGTATTTTCTCCAGCAAGCAAAATCTTTGTTTCAAGTGAACATATTTCTTTT 720
 DB 7736 ACTGTAGAGATTTTCTTTGCAATTAACCAATCTCTTTTGAAGCGAAATATTTAGCTTC 7795
 QY 721 TATACACGCTCTTGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 780
 DB 7796 TATACACGCTCTTGAAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTT 7855
 QY 781 GAGACGAGGTTTGTTTTCAATACCAATCTCTTCCAGCTATGGAACGCTGCTTCTTG 840
 DB 7856 AACACTACTATTTACTTTTAAATGTTGTAACCTTACGCGAGTCAATGAGCGTGCCTGCTT 7915
 QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTAATCAAAATCAATTTCA 900
 DB 7916 TTTCAAGTGGGACTCAAAATGGTACTGTTAAGCTTGAGATTAATCAAAATCAATTTCA 7975
 QY 901 GCTCATGTTAACTCACTGAGTTGGTAAAGTAAAGGAGGATTTAGATATTTGTTAGTCA 960
 DB 7976 GCTCATGTTAACTCACTGAGTTGGTAAAGTAAAGGAGGATTTAGATATTTGTTAGTCA 8035
 QY 961 TCTGTAGTGAATTAATCACTATCAGCTTCAATCCAACTTACCTTTATTTAGTCTTTTAAAGCT 1020
 DB 8036 ACTGTGAGATTTGACCATTTAGTTTCAACCACTTACTTTGATTTGATTTCTTTAAAGCT 8095
 QY 1021 ATTTAAAGTGAACAGTAAATTTCAATTTTATCAGCAGTTTCAACCACTTACTTTTAAAGCT 1080
 DB 8096 TTTAAATAGCAAGAGGAGTAAATTTTATCAGCAGTTTCAACCACTTACTTTTAAAGCT 8155
 QY 1081 CCAGCGGATGAGGAAGTAAATTTTCAATTTTATCAGCAGTTTCAACCACTTACTTTTAAAGCT 1133
 DB 8156 CCAGCGGATGAGGAAGTAAATTTTCAATTTTATCAGCAGTTTCAACCACTTACTTTTAAAGCT 8208

RESULT 24
 BD003700
 LOCUS 21338 bp DNA linear PAT 31-JAN-2002

DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
 ACCESSION BD003700.1 GI:18631661
 VERSION JP 2001501833-A/20.
 KEYWORDS unclassified
 SOURCE unclassified
 ORGANISM unclassified.
 REFERENCE 1 (Bases 1 to 21338)
 AUTHORS Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C., Fannon,M. and Dougherty,B.A.
 TITLE Polynucleotide of Streptococcus pneumoniae and sequence
 JOURNAL Patent: JP 2001501833-A 20 13-FEB-2001;
 HUMAN GENOME SCIENCES INC
 COMMENT OS Unidentified
 PN JP 2001501833-A/20
 PD 13-FEB-2001
 PF 30-OCT-1997 JP 1998520718
 PR 31-OCT-1996 US 60/029960
 PI CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN, PI STEVEN C BARASH,
 PI MICHAEL FANNON,BRIAN A DOUGHERTY
 PC C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,
 PC C12N1/21,
 PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC G06F15/40
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..21338 /organism='Unidentified'.
 FT Location/Qualifiers
 1..21338 /organism='unidentified'
 /mol_type='genomic DNA'
 /db_xref='taxon:32644'

Query Match 53.6%; Score 608.2; DB 6; Length 21338;
 Best Local Similarity 71.1%; Pred. No. 3.2e-102;
 Matches 805; Conservative 0; Mismatches 328; Indels 0; Gaps 0;
 QY 1 ATGATTCATTTTCAATTAATCGACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
 DB 7076 ATGATTCATTTTCAATTAATTAATTAATTTATTTCTACAGCATTAATTAATTAATTAAGAGA 7135
 QY 61 GCTATTAGCACTAAATGCAATTCCTATTCTTTTCATCAATAAAATTTGAAGTCACTTCT 120
 DB 7136 GCTATTAGTCTAAATGCAATTCCTATTTTATCAACAGTAAATTTGACGTGACCAAT 7195
 QY 121 ACAGAGTAACTTTTAAACAGGCTTAAACGGTCAAAATATCAATTTGAAACACATTTTCTGTA 180
 DB 7196 GAAGGTATTAATTAATGTTTCAATTTGTTTCAATTTGTTTCAATTTGAAATTTTATTTCTCAA 7255
 QY 181 AGTAATGAAATGCTGGTTTCTGTAATTAACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 DB 7256 AAAAATGAAGATGCTGGTTTCTGTAATTAATTAATTAATTTAGTTCGATCCCTCTTGAAGCTTCT 7315
 QY 241 TTTTATTAATTAATTTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTGAA 300
 DB 7316 TTTTATTAATTAATTTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTGAA 7375
 QY 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAAAGATGTT 360
 DB 7376 CAAAATCAAAATGTTTAAACAGTGGTAAATCAGAGATTAACCTTAAAGGAAAAAGATGAGC 7435
 QY 361 GACCAAGTATCTCGTCTCAAGAAAGTATCAACAGAAATTCCTTTTGAATTTTAAAGAAACAAA 420
 DB 7436 GAACAAATATCCAGAAATCAAGAAATTTTCAGCAAGCTCTCTTTAATTAATTTGAAACAAA 7495
 QY 421 TTTTGAAGTCTATTAATTTGCTGAAACAGCTTTTTCAGCAGGTTTCAAGAAAGTCTGCTCT 480
 DB 7496 TTTTGAAGTCTATTAATTTGCTGAAACAGCTTTTTCAGCAGGTTTCAAGAAAGTCTGCTCT 7555

[illegible]

[illegible]

```
Db 4567 GCTATTAGTCTTAAATGCGCAATCTCTATTATCACTGTCAAAATGACGTGACCAAC 4626
Qy 121 ACAGGAGTAACCTTTAACGGGTCTAACCGTCAAAATATCAATTTGAAAAACATATCTCTGTA 180
Db 4627 GAAGGTGTTACTTTAATGTTTCAATGCTCAAAATTCATTTGAAAAATTTTATTCTCAA 4686
Qy 181 AGTAATGAATGCTGGTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Db 4687 AAAAATGAAGATGCTGGTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4746
Qy 241 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
Db 4747 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4806
Qy 301 CAACACCAAGTTGTTTAAACCAAGTGGTAATCAGAGATTAATTAATTAATTAATTAATTAAT 360
Db 4807 CAAATCAAAATGTTTAAACCAAGTGGCAATCAGAGATTAATTAATTAATTAATTAATTAAT 4866
Qy 361 GACCAATATCCTCGTCTACCAAGAGTATCAACAGAGAAATCCTTTGATTTTAAACCAAAA 420
Db 4867 GAACCAATATCCTCGTCTACCAAGAGTATCAACAGAGAAATCCTTTGATTTTAAACCAAAA 4926
Qy 421 TTATTGAAGTCTATTATTGCTCAACAGCTTTTGCAGCCAGTTTCAACAGAAAGTCTGCTCT 480
Db 4927 TTACTCAAGAAATTAATTAATGAACAGCTTTTGTGCAAGTACACAGAGAGTCTGTCG 4986
Qy 481 ATTTTAAACAGGAGTTTCAATATGTTTAAAGTAAATCAATAAGATTTTAAAGCAGTAGCGACT 540
Db 4987 ATTTTAAACAGGAGTCTCCTGTTGATTAAGTCAACCAAGAGTAAACCAAGTCTGCAACA 5046
Qy 541 GACTCTCATCGTATGAGCAAGCTTTAATCACTTTGGCAATATCACTTTGAGCAGATTTGATG 600
Db 5047 GACTCTCATCGCTAAGCCAGAGAAATTAATCACTTTGAGCAATATCACTTTGATTTGAT 5106
Qy 601 FTAGTTCTTCAAGTAAATCTTTGAGAGATTTTTCAGCAGATTTTACAGATGATATTGAG 660
Db 5107 GTGCTAATCTTAGCGGTTCTCTACGCGAATTTTTCAGCGGATTTTACAGATGATATCGAA 5166
Qy 661 ACCGTGAGGATTTTCTTCCACCAAGCCAAATCTTTGTTGAGAGTGAACACATTTCTTTT 720
Db 5167 ACTGTAGAGATTTCTTCTGCCAATAACCAATCTCTTTTGAAGCGAAATATTAGCTTC 5226
Qy 721 TATACAGCGCTCTTGAAGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
Db 5227 TATACTCGTCTCTAGAGGAAATCTATCTGTATACAGATCGCTTGTATTTCCACAGACTTT 5286
Qy 781 GAGACGGAGTTGTTTCAATACCCAAATCCCTTCGCCAGCTATCGAAGTCCCTCTTGTG 840
Db 5287 AACCTACTATTACTTTTAAATGTTGAATCTTACGCCAGTCAATGGAGCGTCCCGCTTT 5346
Qy 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
Db 5347 TTATCAAGTCCGACTCAAAATGGTACTGTTGAACCTTGAATTAAGATGGGTGTTAGC 5406
Qy 901 GCTCATGTTAACTCACTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTTAGTCAG 960
Db 5407 GCCCATGTTCACTCTCCAGAGTTGGTAAAGTAAACGAGGAAATCGATCACTGATCAGTT 5466
Qy 961 TCTGTTAGTATTAATCACTATCAGCTCAATCAACTTCTTATGAGTCTTTTAAAGCT 1020
Db 5467 ACTGGTGAAGATTTGACCAATTAAGTTTCAACCCCAATCTTCTGATTTCTCTTAAAGCT 5526
Qy 1021 ATTTAAAGTGAACAGTAAAAATTCATTTCTTTATCACCAGTTTCGACCATTTACCCCTAACA 1080
Db 5527 TTAAATAGCGAAAGGTGACCATTAGCTTTATCTCAGCTGTTCTGTCCTTACTCTGTG 5586
Qy 1081 CCAGCGCATGAGGAGAAAGTTTATCCAAATTAATTAACACAGTACGAAACAA 1133
Db 5587 CCAGCAGATCTACGAGAGACTTCATGCACTCATTACACAGTTCGTACAA 5639
```

```
AE008385
LOCUS 14222 bp DNA linear BCT 13-SEP-2001
DEFINITION Streptococcus pneumoniae R6 section 1 of 184 of the complete genome.
ACCESSION AE008385 AE007317
VERSION AE008385.1 GI:15457528
KEYWORDS Streptococcus pneumoniae R6
SOURCE Streptococcus pneumoniae R6
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 14222)
AUTHORS Hoskins, J.A., Alborn, W. Jr., Arnold, J., Blaszcak, L., Burgett, S., DeHoff, B.S., Estrem, S., Fritz, L., Fu, D.-J., Fuller, W., Geringer, C., Gilmour, R., Glass, J.S., Khoja, H., Kraft, A., Lagace, R., LeBlanc, D.J., Lee, L.N., Lefkowitz, E.J., Lu, J., Matsushima, P., McAhren, S., McHenry, M., McLeaster, K., Mundy, C., Nicas, T.I., Norris, F.H., O'Gara, M., Peery, R., Robertson, G.T., Rockey, P., Sun, P.-M., Winkler, M.E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C., Baltz, R.H., Jaskunas, S.R., Rosteck, P.R. Jr., Skatrud, P.L. and Glass, J.I.
TITLE Genome of the bacterium Streptococcus pneumoniae strain R6
JOURNAL J. Bacteriol. 183 (19), 5709-5717 (2001)
MEDLINE 21429245
PUBMED 11544234
REFERENCE 2 (bases 1 to 14222)
AUTHORS Hoskins, J.A., Alborn, W. Jr., Arnold, J., Blaszcak, L., Burgett, S., DeHoff, B.S., Estrem, S., Fritz, L., Fu, D.-J., Fuller, W., Geringer, C., Gilmour, R., Glass, J.S., Hann, A., Khoja, H., Kraft, A., Lagace, R., LeBlanc, D.J., Lee, L.N., Lefkowitz, E.J., Lu, J., Matsushima, P., McAhren, S., McHenry, M., McLeaster, K., Mundy, C., Nicas, T.I., Norris, F.H., O'Gara, M., Peery, R., Robertson, G.T., Rockey, P., Sun, P.-M., Winkler, M.E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C., Baltz, R.H., Jaskunas, S.R., Rosteck, P.R. Jr., Skatrud, P.L. and Glass, J.I.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA
FEATURES
source
1. 14222
/organism="Streptococcus pneumoniae R6"
/mol_type="genomic DNA"
/strain="R6"
/db_xref="taxon:171101"
1. 1362
/gene="dnaA"
/note="synonym: spr0001"
1. 1362
/gene="dnaA"
/codon_start=1
/transl_table=11
/label=spr0001
/product="DNA biosynthesis, initiation, binding protein"
/protein_id="AAK98805.1"
/db_xref="GI:15457529"
/transl_table=11
/label=spr0002
1. 2657
/gene="dnaN"
/note="synonym: spr0002"
1521. 2657
/gene="dnaN"
/EC_number="2.7.7"
/codon_start=1
/transl_table=11
/label=spr0002
```

```

/product="DNA biosynthesis; sliding clamp subunit,
required for high processivity; DNA polymerase III
beta-subunit"
/db_xref="GI:15457530"
/translation="MHFSINKNLFQALNITKRAISSKNAIPILSTVKIDVNEGVT
LIGNSGQISFENFISQKDEAGLITSLAGSILLESFFINVSSLPDVLDFKEIBON
QIVTSKSEITLQKQSEQYPRIQISASTPLILETKLKKIINETAFAASTQESRP
ILTGVPVLSQKELKTVADSHRLSKKLTLEKNSDDPDVVPSPSLREFAFTDD
IETVEFPFANQILFSENISFVTRLEGNYPDTDLIPDTENTITFNNVNAQSMDE
RALLSATQNGVKLEIKDCVSAHVHGSPEVKGNEEDITDOVGTEDLITISFNPTYL
IDSIALNSEKVIISFISAVRPFTLVPADTDEDFMQLITPVRTN"
2722..2916
/gene="spr0003"
/label="spr0003"
/translation="MYQVGNFVEMKPHACTIKSTGKKAARWEITRVGADIKKCSNC
EHVVMGRYDFEKKMKIID"
3000..4115
/gene="spr0004"
/label="spr0004"
/translation="WALTAGIVGLPNVCKSTLFLNATKAGABAANYPPATIDPNVGMV
EYDRLQKLTWMTPKKTIPTTFEFTDIAGLVKAGSGEGGNKFLANIREVDAIVH
VRAFDENMREQREDAFVDLADITINLELILADLESVKNRYARVERKARTQKD
KESVAFNQLKIVLEDGKSARTIEFTDEQKVVGKGLFLITKPVLYVANVEDDV
SPDSIDYVQIREFAAENAENVVISARAEIISLDEDEKKEFLEALGTESGVDK
LTRAAYHLIGLGTFTAGEKEVRWTKRGMKAPQAGI IHSDPEKGFIRATVTSYED
LVYGSKAVKEAGRLREBEGKEYIVQDGDIMEFRNV"
4186..4755
/gene="pth"
/label="pth"
/translation="MTLLVGLGNPGDKVFKTNKVGPMILIDOLAKQNVTFTHDKIF
QADLASFLNGEKIYLVPTPTFMESGKAVHALLTYGLDIDDLIIYDDLDMEVGKI
RLRAGSGAGNGIKSIIOHIGTQVFNKVIKIGIRPKNGMSVHHVLSKFRDDYIGI
LQSIDKDVDSVNYILOEKKFKTQRYNG"
4756..8265
/gene="mfd"
/label="mfd"
/translation="MWTLDLFSENDQIKKHONLTDKKRQLILGLSTKALAIASS
LEKDRILVTSYGEAEGLVSDLSILGELVYFPFLVDDAPMVFFLMSQEKIISRV
BAKRLFDSSKGLVLCNIAASRLILPSNAPKDSIVKISVGEYDQHAFTHOLKNG
YRKVTQVOTQGEFSLRGDILDFEISOLEPCRIEIFFEGDEIDGIRSFVEVQLSKENT

```

```

ELTIPASDMLLRKDYORQGSALFKQISKTLSPILKSYLBEILSSFHQKSHADSRK
FSLCYDKTWIVFDYIEKDTPIFFDDYQKLMNQYEVFERELAQYFTEELQNSKFSDM
QYFSDIEQIYKQSEVTFPSNLQKGLNKFQKIYQFNQYPMQEFNQSFLKEIER
YKMDYITILQSSNSMGSKTLEDLMLEQIYKIDSRDKTSICKESVNLTEGLRHGFHF
VDEIKILLTEHEIFOKKLRFRROHVSNAERLKDYNLEKGDVYVHHIHGIGOVGLI
ETIEKETHRVDSVOYONGQOISIPVQIHLISKYISSDGAKPKLKLNDGHFKKAK
KQKNOVEDIADDLIKLYSERSQLGFASADDDDDQDAFFPVTETDDQLRSLEEI
KRMQASQPMRLLVGVDPGKTEVAMPAFAKVNHDHQQVILVPTTVLAQOHTNFK
ERFQNFANVNLVSFRSKQBTATLEKLNQGVQDILIGTHRVLSKDVVFDLGLMI
DEQRFVGHKRTIKELKQVQDVLTLATPIPTLHMSMLGIROLSVIETPTNRPV
QTYVLEKQNSVIRDAVLREMERGGQYVLYNKVDITVOKVSELOELIPEASIGVYVHG
MSEVOLENTLDFIEGOYDILVTITITETGVDIPNANTLFIENADHMLSLYLQLRG
VGRNRIAYAVLMYRPEKSISEVSEKLEAIKGTLSGSGFKIAMRDLISKAGNLLG
KSQSPIDSVGFELYSOLLEAIAKRNANANTKGNAEILIQIDAYDPTTISDQ
RHKEIYKRIQIDNRVNEBELIDRFGEYDPVWAYLEILGLVSKYDKVFORV
ERKONKITQIEKVTORLFQADYFKALSVTNLKAGIENKGLMELVDFVQNKDYEI
LEGLIFGESLLEIKESKEKNSI"
8323..8589
/gene="spr0007"
/label="spr0007"
/product="Conserved hypothetical protein"
/protein_id="AAK98811.1"
/db_xref="GI:15457535"
/translation="MRLDKYLKVSRIIRKRTVAKEVADKGRIKVNGILAKSSTDLDKN
DQVEIRFGNKILLVKLEMKDKSTKKDAAGMYEIISETRVEENV"
8582..8950
/gene="spr0008"
/label="spr0008"
/product="Hypothetical protein"
/protein_id="AAK98812.1"
/db_xref="GI:15457536"
/translation="MSKNIVQLANSFIONERYORRYLMKROKRNPMGVLILMLL
FIUPTFNLAQSYQQLLQRRQOLADLQYQTLSDKDKETAFATKLKDEYAAKYTRA
KYYSKSRKVTIPDLLQR"
9001..10338
/gene="spr0009"
/label="spr0009"
/product="Conserved hypothetical protein"
/protein_id="AAK98813.1"
/db_xref="GI:15457537"
/translation="MKSWMNWLKINVCYVKRKRGRMRKFLIILLPSFLTITSKVS
TEKEVYSKEIYLSQSDPGFYFREKLSPPMYGVYVYANEDLVVSGKLTPKTSF
QITWEKLNQGIPIVFKLSHQFIADKRELYDQSEVTPITIKVWLESDFKLYNSPYDL

```

```

Query Match      53.2%; Score 603.4; DB 1; Length 1422;
Best Local Similarity 70.8%; Pred. No. 2.7e-101;
Matches 802; Conservative 0; Mismatches 331; Indels 0; Gaps 0;

Qy 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATACAACTAAACGT 60
Db 1521 ATGATTCATTTTCAATTAATTAATAAATTTATTTCTACAGCATTAATATTACTAAGA 1580

Qy 61 GCTATTAGCACTAAAATGCCATTCCTATCTTTTCATCAATAAAAATGAAGTCATTCT 120
Db 1581 GCTATTAGTTCTAAAATGCCATTCCTATTTATTAACAAGTGTCAAAATGACGTGACCAAC 1640

Qy 121 ACAGGAGTAACTTTAAACAGGCTTAACCGGTCAATATCAATCAATGAAACACATTTCTCTGA 180
Db 1641 GAAGGTGTTACTTTAATTTGGTTCATATGTCATATTTCAATGAAAATTTTATTTCTCAA 1700

Qy 181 AGTAATGAAAATGCTGGTTGCTAATTAATCCTCTCCAGGAGCTATTTATTTAGAAAGCTAGT 240

```


Qy	541	GACTCTCATGTTATGAGCGAAAGTTTAAATCACTTTGGACAAATACCTTACAGACAGATTTGATG	600
Db	2422	GACTCATCGGATGATGATCAACCGCTTTATTGGCTTTTAAAGAGATTCAGAAATTTAAATTTGAT	2481
Qy	601	GTTAGTTCTTCCAAAGTAAATCTTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATTTGAG	660
Db	2482	GTTATTTTGGCAAGTAAATCAATCAATAGTTTAAATAATGTTTTCACAAATGATGAAGAA	2541
Qy	661	ACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAAAGTGAACACATTTCTTTTT	720
Db	2542	GAGATTTGAAATTTTCATTAGTGAAGCCAAATGCTTTTATAGAAATGAACAATAGTTAT	2601
Qy	721	TATACAGGCTCTTAGAAGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT	780
Db	2602	TATAGTCGTTTGAATTTGAAGGATCTTATCTGTACTAAATCGTTTAAATTCACAAATGAAGCT	2661
Qy	781	GA-----GACGGAGGTGTTTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTC	834
Db	2662	GATTATCTTTAGATTTAGTTTGTGCTGCGCAATTTGGCTATCAATGGATCGTGCG	2721
Qy	835	TTCTTGATTTCTAATGCTACTCAAAATGGTACTGTTAAAGCTTGAGATTTACTCAAAATCAT	894
Db	2722	CGTTTGTGACAGTAATGACAACATAATGGAACCTGTAATAATTAACGGTTTCTGGTGATCT	2781
Qy	895	ATTTACAGCTCATGTTAACTCACTCGAGGTTGGTAAGGTAAACGAGGATTTAGATATTGTT	954
Db	2782	GTAGTAACACACAGCCAACTCTCTGAGGTAGGTTCTGTCCATGAAGAATTTGACGGCCTTA	2841
Qy	955	AGTCAGTCTGGTAGTGCATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTA	1014
Db	2842	TCAAAAGAGGAATGACTTAGCAATCAGCTTTAATCCCTGAATTTTGAATGATGCTTTG	2901
Qy	1015	AAAGCTATTAAAAAGTGAACACAGTAAAAAATTCATTTCTTATCACCAGTTTCACCATTCACC	1074
Db	2902	AAAGTGATTAAAGCACACAGAGTTCGTATTCGGTTTATTTCTAATGTTTCGTCGGTTCACT	2961
Qy	1075	CTAACACCGGCGATGAGGAAGAAAGTTTATCCAAATTAATACACAGTACGAAACAAA	1133
Db	2962	TTACAGCCACGTAAATGAAGAACTCGGATTTGTTCAATTGATTACACAGTACGTACAAA	3020
RESULT 31			
LOCUS	LLU76424		
DEFINITION	Lactococcus lactis DnaA (dnaA) gene, partial cds; polymerase III subunit DnaN (dnan), exonuclease RexB (rexB), and exonuclease RexA (rexA) genes, complete cds; putative GTP binding protein gene, partial cds; and unknown gene.		
ACCESSION	LLU76424		
VERSION	U76424.1		
KEYWORDS	GI:2909712		
SOURCE	Lactococcus lactis		
ORGANISM	Lactococcus lactis		
REFERENCE	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.		
AUTHORS	1. (bases 1 to 10762)		
TITLE	el Karoui, M., Ehrlich, D. and Gruss, A.		
JOURNAL	Identification of the lactococcal exonuclease/recombinase and its modulation by the putative Chi sequence		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 95 (2), 626-631 (1998)		
PUBMED	98118563		
REFERENCE	2. (bases 1 to 10762)		
AUTHORS	El-Karoui, M. and Gruss, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-OCT-1996) Genetique Microbienne, INRA, Domaine de Vilvert, Jouy-en-Josas 78356, France		
FEATURES	Location/Qualifiers		
source	1. .10762		
	/organism="Lactococcus lactis"		
	/mol_type="genomic DNA"		
	/strain="WG1363"		
	/db_xref="taxon:1358"		

gene

1. .820

/gene="dnaa"

<1. .820

/gene="dnaa"

/function="initiation of DNA replication"

/codon_start=2

/transl_table=11

/product="DnaA"

/protein_id="AAC12963.1"

/db_xref="GI:2909713"

/translation="IKVSSNFVNDVYNAIYTRNQMESFNTYRNLDLDDVQFFS
 DKEGTRNEFFNFALYDKGQIVLTSRDIPOQLNLEDLVSRFGWGLTDTTADPY
 ETRWAILLISBSHFPSETLSYTAGQIDSNVRELGALNVEFVARANGISIVDI
 YTAQALRSIKNAATQOGLSNLTIKKIODEVANYTHISFSLVGPKEPAEAPROIAM
 YLRELCTSLPAGTGTAGCRDHTTVMVYKQISDRKMDNDVQKDIDSIKKPF"

977. .2119

/gene="dnaa"

977. .2119

/gene="dnaa"

/codon_start=1

/transl_table=11

/product="polymerase III subunit DnaN"

/protein_id="AAC12964.1"

/db_xref="GI:2909714"

/translation="MKKFSINKTAFQNALIKTKQAATGSKVTIPALTKLIEVEERGIT
 LIGSGQISIKNLPADNDKASMLISGTGSVLEAFNFVVSQLPVLEPIEKEQK
 QVLTSGKSTIKLGLDSEIYPHLOISSEGSLKMKVKLEIFETFTFAVSTQENRP
 IFVHLETLSTGELKAVATDSHRMSQRLPLEDTDLKFDVILPSKINSKPNVFTND
 EEFIEIFISGQMLQONETISYYSRLIEGYSPTNRLINEADYTLDLVFDNAQURHT
 MDRALILVTMTNGTVKLVSGDSVVTANSPEVGSVHEELTALSKGENDLSISFNP
 YLTDALKVKAPVVRIRFISNVRPFTLPQRNEESGFVQLITPVRTN"

2225. .5524

/gene="rexB"

2225. .5524

/gene="rexB"

/note="second subunit of major exonuclease"

/codon_start=1

/transl_table=11

/product="exonuclease RxB"

/protein_id="AAC12965.1"

/db_xref="GI:2909715"

/translation="MELIYTEITQDLTEGLLEIALELEKKNRVYIIVPSSWSPEKEK
 EILERLAGSDTAVFDLIVTRKQPPYFDKREKATMTGLTGLSLMFLRVLRSFK
 KDBIPLYFSLDSAGFLEMLIOLEAELTANLISVENPDNPNKQELKILAKFPAELS
 VEVANYSEGDFTNLDVGEFQDLKDVITIIDGTRFSAEELFIESIQKVARFV
 GTYSNLSLTASEIYVGTSMITFRNKPFPVELRKAASAVNEVSKLRILDLDS
 RVITIDEKLEAEDEKFRIVAEQNEKVEIRVAKETIRQKIIQCAFPPKDFVLVGP
 AAYEITLKEVFLYIEIPFYAQEESMSQHLPIVFESLFAIKKNRYTDDVNLKSK
 VYTDANLDEVIDYFEYVYQYKISGRKKFTBEFSEFSOIELVNEMRKLLGSESP
 LQVFLNNRKTGKKWSDLOGLLENGNVMTNMYAPSAELQNEHQWAKHEQVWQM
 LISTNEFLAVPSDEKLSVFLDILLAGLAKAKYRQIPANVDVNVKDYELVEPFTM
 KYIYALGHSQTNFPIKKNSTLSDDEERLEINTDENQFIEQLNVANYQKQFTVLS
 LINSKESLIVSMPOIMANEQEFSPVQLFKDADEKILQIKQVNLFSLEHIGNS
 RSVIANTQIKRELVESEDEKRVFWSIFRILVKSNDAPFKILLDLAKDIDTVNL
 APDLTEUIGDKIYASVSSFERFYCYEPLNTLSLETENIDNSKLVGNFFHEV
 FAKVMTEDLAENFDEKLTLOVDKNVSRFTQDARFTWNSLEEVIRGTATVL
 KATVSTDLKTLITLESSPLPKSELGNFSDVDIYLRGIDELDOLSDYDLCADYKSS
 AHSFKLOAYDGLSLQFTYLDVILKQAPPNQKINGALYLOFKQPINLSINQUSIET
 NILKESMEYEGVLDEAAEQIKGIENTALKNTINYNEEFELKLKNEEHYRAAGQRL
 KKGKAIINPMKREGIDQSGNVGRGCRYCPUKSICRPFANIHMHNEHSREIGQKSOAEI
 LAELGEERDE"

5517. .9038

/gene="rexX"

5517. .9038

/gene="rexX"

/note="first subunit of major exonuclease"

/codon_start=1

/transl_table=11

/product="exonuclease RxB"

/protein_id="AAC12966.1"

/db_xref="GI:2909716"

/translation="MSEVKLTPEQNEAIHSSGKNILVSAAGSGKTFVMAQRIVEKVK
 QGIEIDLFIISTFTKKAASELRMLERDLAKARQESSDDEEHRLLTALQNLNADIG

TMSFTQKLTKANFNVRNIDPNFRILADQTESDLIROBVSQVLESYLSADESNLISK
 DKFEKLIKNSKDRNIILOKVVVTVYFASATENPISLWENQELKGETYSAESNLISK
 EDFTVNVKENILLTREPILLEALSKKDPVCTALFLSIDDIDRVGSSKDALSKKKDFS
 AQKDLVGSKSPGELRFVDKIKHGDIETKIQNAFEIASDLQKFIIDFYKTYLERK
 KNEAFESDIATFAIEILEENPDIRENLREHYDEIMIDEYQDTSHTQERMLELLSNG
 HNLFWGDIKOSIYFGRADPLGLEKYKSYDOAKENQNLIRLKENFRSGEVLNFTN
 DIFHLMDEKLGEMTYGKEALVQGNISDYPVEAKQFPELLYKKNLFRSSEETEDSE
 VIKSDGEIKGAAQIKKLIYGVPEPKDIAILVRSKNNKTEDILLSDVIPVLDGR
 VDFLKSMEVIMLDVLRADNPLVDLSIVMLRSPLPFNEDELTTRISVQGSRODRFW
 DKLLSLKEGKNPELINLSLEQKLFKAFNQKFTWRKLVNKPILHRLLLKTYITYYP
 DYVGLKNGENRQANQLSVRAESYESSGYKGLFKFVRLINKMEQNNDIASVNIKL
 PQNAVTRMTFHSKLEGFYVFLMNLQSFNDRDLKEDVILSRHGKGLYADLKA
 DDVTFPYALVAKMETTFYVMWKDLKQRAALSEEMRVLYAFTRAKKLYLVGKIKOT
 DKVAGLELDAATLEGLKILSDKFRNSRGFOHWILALQNAATKLPKMLNVTKDLETE
 KLEFSTQDPFKKLVVESEKFDNISFSDSEIKEAOKIWNQYPHQAATLSSIQTPSQV
 KKRSEYQLOVGEVQVSEFVRVKNLDFSDGPKKITAENGSAETHSFMQYADFPQD
 LFSQATLDEMGPDEKLNQIDITKILITLFDTEFGQFLSENVDKTVKEAPFSMLRTD
 FAEQYIVRGICDGFVKLADKILIDFYKTRFTNVSAISEIKERYKQDMNLYSEALQK
 AYHVNQDKLIVLILGGRKPVFEKIDD"

CDS

complement (9281. .9838)

/note="putative ORF 5"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC12967.1"

/db_xref="GI:2909717"

/translation="MLGENIQKARLSKNLTQEEVAKELYFSROAISRWESNKTENPFE
 TLALADLYSDLLAFSQGIEPOKRKKHINLPAGFGLVPFNFLVGLFWLLIITALLLLA
 LYGLILAIPLAFIMIPALVTHNPHTTFGNANGATGDMWYWLVAIISCSLMVIVSFPF
 IWRFTKLLYFWLLKYLKFNLSVYN"

CDS

10122. .>10762

/note="ORF 6"

/codon_start=1

/transl_table=11

/product="putative GTP binding protein"

/protein_id="AAC12968.1"

/db_xref="GI:2909718"

/translation="MALTAGIVGLPNVKGSTLFPNAITKAGABAAANPPATIDPNVGMV
 EVDPELNKLTTELKPKKTVPTTFEFTDIAGIVKGSAGEGLGNFKLANIREVDIAIH
 VVRAFDENVRNRENNEDAFIDPMADIENTINLELILADLESVYKRYARVEKARTAKD
 KDAVAFNVLKLLKVLDEGKSARTIDFDEDETKVLKSLFLLTSKPVLYVANVS"

ORIGIN

Query Match

Best Local Similarity 36.0%; Score 408.2; DB 1; Length 10762;

Matches 690; Conservativity 0; Mismatches 443; Indels 6; Gaps 1;

Qy	1	ATGATTCAATTTTCAATTAATCGCATTATTTATTTATTCATGCTTTAAATAACAACCTTAAACCT	60
Db	977	ATGATTAAATTTTCAATTAATAAATGCGCTTTCAAAATGCTTTGAAATTTACCAACAA	1036
Qy	61	GCTATTAGCACTAAATAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGCTACTTCT	120
Db	1037	GCAATTTGGTTCTAAGTTACAATTCCTGCTTTTAAACAAAGTTAAAAATTTGAAGTTGAAGAA	1096
Qy	121	ACAGGAGTAACTTTAAACAGGGCTAACCGCTCAAAATCAATCAATGAAACCACTATTCTCTGT	180
Db	1097	AAAGGAATTAACCTTAATTTGTTTCAATGGTCAAAATTTCAATCAAAAACCTTTTTCAGCT	1156
Qy	181	AGTAATGAAATTCCTGGTTTGTCTAATTAACCTCTCAGGAGCTATTTTATTTAGAAGCTAGT	240
Db	1157	GATAATAAGATGCAAGTATGCTTTATTTTTCAGAACTGTTCTGTCTTATTAGAAGCTGCT	1216
Qy	241	TTTTTTTATTAATTAATTAATTTCAAGTTTTCAGATATTAGTATAAATGTTTAAAGAAATTCAA	300
Db	1217	TTCTTTTGAAGATAGTAGTAAGTCACTTCTGAGTTACTTCTGGAATTTATCGAAAAAGAA	1276
Qy	301	CAACACCAAGTTGTTTTTAAACAGTGGTAAATCAGAGATTACTTTAAAGAGAAAGATGTT	360
Db	1277	CAAAAACCAAGTTTGTGTTGACTTCTGGTAAATCAGAAATTAATTTTAAAGGTTTGAATTC	1336
Qy	361	GACCAAGTATCTCGTCTACAAGAGATATCAACAGAAATCCCTTTGATTTTAAAAACAAA	420

Db 1337 GAAATTTATCCACATTTCAAGAAATCTCTGAAGGTTCTTCTTAAATGAAAGTGAAA 1396

QY 421 TTATTTAGAGTCTATTATTTCTGTAAGAGCTTTTTCAGCAGGTTTCAAGAAAGTCTGCTCT 480

Db 1397 GTTTTGAAGAAATTTTACAGAGAGTCTGTTTTCGTGTAAGTACTCAAGAAATGCTGCA 1456

QY 481 ATTTTAAAGAGGAGTCTATATTTCTATTAAGTATATCAATGAAGATTTTAAAGCAGTACGACT 540

Db 1457 ATTTTCAAGAGGAGTCTATTAAGAAATTTTGTCAAGGGTGAATTTAAAGCAGTACGACT 1516

QY 541 GACTCTCATCGTATGAGCAAGCTTTTAAATCACTTTGGCAATATCTTACGACAGATTTTGATG 600

Db 1517 GACTCATCGAATGAGCAAGCTTTTATGCTTTTAGAGATACAGACTTAAATTTGAT 1576

QY 601 GTAGTTCTTCCAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATATGAG 660

Db 1577 GTAAATTTTACCAAGTAATCAATCAATAGTATTTTAAAGATGTTTTCACAAATGATGAAGAA 1636

QY 661 ACCGTTGAGGATTTTCTCACCAGCAAGCTTTGTTTTCAGAGTGAACACATCTTCTTTT 720

Db 1637 GAGATTGAAATTTTCACTAGTGAAGCAAGCTTTTTCACAAATGAAACAAATGATGTTAT 1696

QY 721 TATACACGCTCTTGAAGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780

Db 1697 TATAGTCTGTTGATGAGGTTCTTATCTGATATCTATCTGTTTGAATCCAAATGAGCT 1756

QY 781 GA-----GACGAGGAGTTGTTTCAATACCCAAATCCCTTCGCCAGCTATGGAACGTGCC 834

Db 1757 GATTATATCTTTAGATTTGCTTTTGTGCTGCGCAATTCGCTCATCAATGATGCTGCA 1816

QY 835 TTCCTGATTTCTAAGTCTACTCAAAATGCTACTGTTTAACTTGAAGTCTTCAAAATCAT 894

Db 1817 CGTTTATTGACAGTAAATCAACCAAGCAAGTAAATGCTGTTTCAAGTGAATCT 1876

QY 895 ATTTTCAAGCTCATGTTTAACTCACTGAGGTTGCTGAGTAAAGTAAAGAGGATTTAGATTTGTT 954

Db 1877 GTCTGTAACACAGCAAACTCTCTGAGTAGGTTCTGTTTCAAGAGACTGACAGCTTA 1936

QY 955 AGTCACTCTGTTGATGATTAATCACTACGCTTCAATCCAACTTACCTTATGAGTCTTTTA 1014

Db 1937 TCAAAAGAGGAGTAAATGATTTATCTATCATGTTTCAATCTGAGTATTTGATGCTTTG 1996

QY 1015 AAAGCTATTTAAAGTGAACAGTAAATTTTATCTATCACTGATTTTCACTGATTTTCACT 1074

Db 1997 AAAGTGAATTAAGACCAAGTAAATTTGATTTTCACTTCACTGATTTTCACTGATTTTCACT 2056

QY 1075 CTAACACAGGAGGATGAGGAGAAAGTATTTTATCCAAATTAATTTACACCACTGACGACAAA 1133

Db 2057 TTGCAACACGGAATGAAGAACTGGAATTTGTTTCAATGATTTACACCACTGACGACAAA 2115

RESULT 32

LOCUS AE016947 326434 bp DNA linear BCT 28-MAR-2003

DEFINITION Enterococcus faecalis V583, section 1 of 11 of the complete genome.

ACCESSION AE016947 AE016830

VERSION AE016947.1 GI:29342119

KEYWORDS

SOURCE Enterococcus faecalis V583

ORGANISM Enterococcus faecalis V583

REFERENCE 1 (bases 1 to 326434)

AUTHORS Paulsen,I., Banerjee,L., Myers,G.S.A., Nelson,K.E., Seshadri,R., Read,T.D., Fouts,D.E., Eisen,J.A., Gill,S.R., Heidelberg,J.F., Tettelin,H., Dodson,R.J., Umayam,L., Brinkac,L., Beanan,M., Dougherty,S., DeBoy,R.T., Durkin,S., Kolonay,J., Madupu,R., Nelson,W., Vamathevan,J., Tran,B., Upton,J., Hansen,T., Shetty,J., Khouri,H., Utterback,T., Radune,D., Ketchum,K.A., Dougherty,B.A. and Fraser,C.M.

TITLE Role of Mobile DNA in the Evolution of Vancomycin-Resistant Enterococcus faecalis

JOURNAL Science 299 (5615), 2071-2074 (2003)

PUBMED REFERENCE AUTHORS

12663927 2 (bases 1 to 326434)

Paulsen,I., Banerjee,L., Myers,G.S.A., Nelson,K.E., Seshadri,R., Read,T.D., Fouts,D.E., Eisen,J.A., Gill,S.R., Heidelberg,J.F., Tettelin,H., Dodson,R.J., Umayam,L., Brinkac,L., Beanan,M., Dougherty,S., DeBoy,R.T., Durkin,S., Kolonay,J., Madupu,R., Nelson,W., Vamathevan,J., Tran,B., Upton,J., Hansen,T., Shetty,J., Khouri,H., Utterback,T., Radune,D., Ketchum,K.A., Dougherty,B.A. and Fraser,C.M.

DIRECT SUBMISSION

TITLE Submitted (03-FEB-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

JOURNAL

FEATURES

source

1.326434

/organism="Enterococcus faecalis V583"

/mol_type="genomic DNA"

/strain="V583"

/db_xref="taxon:226185"

59..1402

/gene="dnaA"

/locus_tag="EF0001"

59..1402

/gene="dnaA"

/locus_tag="EF0001"

/note="identified by match to PFAM protein family HMM PF01695"

/codon_start=1

/transl_table=11

/product="chromosomal replication initiator protein DnaA"

/protein_id="AAO79886.1"

/db_xref="GI:29342120"

/translation="MPDVESFWHSLEBAYQAILSAPSFDAWIKTRPLKLDNNQWLVE VPSAVHRDYWEKNLSAKIVETGKLTGAEMPHVVADEKAAALAELEPAAEEVVF SEQKKAMLNPKYTFDTEVIQKQMAHAALVAEDPGSIYNPLFFGGVGLKTHL MHAICHOMLPQDAKVKYVSEFTNEFINSTQTKSEQRKVRNVDLLVDIQQF LAEATLEEFPHFTNDLYNENKQIVLTSDDPNDI PKLPERLVRFAWGLSVDTIPP DLEVTALIRKKADEARLEIPDDTSLVIAGDINSIRLEGALVRVQAPATNGEDIT TSLAADALSKSVGSKNQLSLQIQEVSKYHVPVLDKLGKRVKVIIVPRQISMY LAREMTDNLSPKIGAEFGGKHDTTIIHAHBIQQLLEKDPATIQKEVSEIKNULNS"

1604..2734

/gene="dnaA"

/locus_tag="EF0002"

1604..2734

/gene="dnaA"

/locus_tag="EF0002"

/note="similar to GB:X69151, GB:J05682, SP:P21283, PID:340188, and PID:37643; identified by sequence similarity: putative"

/codon_start=1

/transl_table=11

/product="DNA polymerase III, beta subunit"

/protein_id="AAO79887.1"

/db_xref="GI:29342121"

/translation="MKLTVRKSVFLQELQTVQRAISSKTTTIPILTVKVLSEGLSL TGSNADISIESFLSKDDKCAQMTERTGSIIVLQSFEGEIIKLPEDMTMEVLNNQ VAITSGADFTVGLDADNPHLPVIDTQNMKLPVHLTKIISTGFAVSHSRPI LTVGHFIDENQKLLAVATDSHRLSORVPTQEAQVEDFNIVPGKLTSLSTNEBE MVEISIMENQVLFKTFMYFSRLGNYPTDRIPTSHNTQIEFYVPELSAIERA SLISGHEGNNIVKLSISPDVLYGNSPEIGKVEALNENVSGEALDISFNPYMKD ALRAFGDNNITVKLSPIRPTLEPETELDIQLITPVRTN"

3313..3558

/locus_tag="EF0003"

3313..3558

/locus_tag="EF0003"

/note="similar to SP:P05650; identified by sequence similarity: putative"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AAO79888.1"

/db_xref="GI:29342122"

/translation="MKKTFVLATEVITIGQFLKEINVIGSGGQAKWYLNADNSVFDVGE LENRGRKLYAGMIEIEEGTFFPMVKNESDDAAE"

gene

CDS

gene

CDS

Db 2081 ATTTTAACTGGGTTCACTTTATTTTAAAGAAATCAAAA---TTACTTGGCCGTCGGACA 2137
QY 541 GACTCTCATCGPATGAGCAACGTTTAACTTTTGGACAATPACTTCAGCAGATTTGATG 600
Db 2138 GATTACATCGTTTAACTCAACGTTGATCCGACAGAACAGCAGTAGAGACCTTTAAC 2197
QY 601 GTAGTCTTCCAGTAATATCTTTGAGAGAAATTTTTCAGCAGATTTTACAGATGATATTGAG 660
Db 2198 ATTTGAAATTCAGGAAAAAGTTTAACTGAACCTTTCTCGTTTCAATTAACCAATGAAGAAGAA 2257
QY 661 ACCGTTTGAGGTATTTTCTCACCAAGCCAAATCTTTGTTTCAAGAGTGAACACATTTCTTTT 720
Db 2258 ATGGTTGAAATCAGCATTTATGAAATCAAGTGCTATTTTAAACAGAAACATGTACTTC 2317
QY 721 TATACACGCTCTTGAAGAAATATCCCGATACAGACCGTTTATTAATGACAGAAATTT 780
Db 2318 TATTCTCGTTTGTAGAGAAATTTATCTGTATACCAACCGTTTAAATCCAACTAGCCAT 2377
QY 781 GAGACGAGGTGTTTTCATACCAATCCCTTCCGACGCTATGGAAGCTGCTTCTTG 840
Db 2378 AACACGCAATTTGAATTTTATGTACCAAGATTTGCTTTCAGCAATCGAAGCTGCTCTTTA 2437
QY 841 ATTTCTAATGCTACTCAAAATCGTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
Db 2438 CTTTCTCATGAGGAGGTAAACATATTTGTCGCTTTCAATTTGACAGATTTCTGTTGT 2497
QY 901 GCTCATGTTAACTACCTGAGGTTGTTAAGTAAACGAGATTTAGATATTTAGTCAG 960
Db 2498 TTATATGAAATTTCACTGAAATTTGAAAGTGAAGAGCTTTTAACTATGAAATGTT 2557
QY 961 TCTGTAGTATTTAACTATACGTTTCAATCAACTTACCTTATTTGAGTCTTTAAAGCT 1020
Db 2558 TCTGTGAGGCTTGATATTTCTTCAACCCAGATTTATGAAAGATGCGTTGCGAGCG 2617
QY 1021 ATTTAAAGTGAACAGTAGTAAATTTCAATTTCTTATCACAGTTTCGACCATTTACCCCTAACA 1080
Db 2618 TTGCGGATGCAATATTTACCGTGAATTTCTTCTCCATTCGTCGCTTACATTTGAG 2677
QY 1081 CCAGCGGATGAGGAAGAAAGTTTATCCAAATTAATTAACACAGTACGAAACAA 1133
Db 2678 CCAACCGAAACAGAACTAGATTTTCACTCAACTTAATTAACCGGTTTCGTACAA 2730

RESULT 33

BD194028 8001 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Enterococcus faecialis polynucleotides and polypeptides.
ACCESSION BD194028
VERSION BD194028.1 GI:33003767
KEYWORDS JP 2002529046-A/622.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 8001)
AUTHORS Kunsch,C.A., Dillon,P.J. and Barash,S.C.
TITLE Enterococcus faecialis polynucleotides and polypeptides
JOURNAL Patent: JP 2002529046-A 622 03-SEP-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Unidentified
PN JP 2002529046-A/622
PD 03-SEP-2002
PF 04-MAY-1998 JP 1998548302
PR 06-MAY-1997 US 60/044031.16-MAY-1997 US 60/046655 PR
PI 14-NOV-1997 US 60/066009
PI CHARLES A KUNSCH,PATRICK J DILLON,STEVEN C BARASH PC
C12N15/31,C07K14/315,C07K16/12,C12Q1/68
CC Strandedness: Double;
CC Topology: Linear;
CC Enterococcus faecialis polynucleotides and polypeptides FH
Key Location/Qualifiers
FT source 1..8001
FT Location/Qualifiers /organism='Unidentified'.
FEATURES

source

1..8001

/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 29.2%; Score 330.6; DB 6; Length 8001;
Best Local Similarity 56.9%; Pred. No. 4e-51;
Matches 645; Conservative 0; Mismatches 485; Indels 4; Gaps 2;
QY 1 ATGATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 598 ATTAAGAATTAACCGTCNAAGGAAGTGTCTTTTACAGAATTTACAACTGTTCAACA 657
QY 61 GCTATTAGCACTAAAAATGCCATTCCTATTTCTTT-CATCAATPAAAAATTTGAAGTCACTTC 119
Db 658 GCAATTTCTTCTAAAAACAACGATTCCTCAATCTTAAACAGGTGTTAAAAATGTTATCAGA 717
QY 120 TACAGGAGTAATTTTAAACAGGCTTAAACGGTCAATATATCAATTTGAAACACTATTCCTGT 179
Db 718 AGATGGCTTATCAGCTACTCTGGGAGTAAACGGGATATTTCAATTTGAAAGTTTAAAGTAA 777
QY 180 AAGTAATGAAAAATGCTGTTTCTTAATTAACCTCTCCAGGAGCTATTTTATTAGAAAGTAG 239
Db 778 AGATGATGAAAAAGCTCAAAATGACCAATTTGAACGACAGGTAGCATCGTTTACAACTCG 837
QY 240 TTTTATTAATTAATTTTCAAGTTTCCAGATATTTAGTATPAAATGTTTAAAGAAATGA 299
Db 838 TTTCTTTGGCGAAATTTATTCGTAATTTACCAGAGATATGTTTCACAATGGAAGTTCTAGA 897
QY 300 ACAACACCAAGTTGTTTAAACAGGTTGTAATCAGAGATTTACCTTTAAAGGAAAGATGT 359
Db 898 TAACAATCAAGTTGCAATTTACTTTCAGGAAAGCTGATTTTACGGTTAATGGAATTAGATGC 957
QY 360 TGACCAAGTATCTCGTCTACAGAGATATCAACAGAAAAATCCTTTGATTTTAAAAACAAA 419
Db 958 TGATAATTTACGCACTTACAGTAATTTGATCTCAAAACCAATGAAATTTACCTGTTCA 1017
QY 420 ATATTGAGTCTATTTATTTGCTGAAACAGCTTTTTCAGCCAGTTTTCAGAGAAAGTCGTC 479
Db 1018 TTTATTGACAAAAATTTATTTAGTGAACAGGTTTTTCTGTCGTATCGATGACGAAAGTCGTC 1077
QY 480 TATTTAACAGAGTTTCAATTTGATTAATTAATCATATAAGATTTTAAAGCAGTAGCGAC 539
Db 1078 AATTTTAACTGGGTTTCACTTTTATTTAGAAAAATCAAAAA---TTACTTGGCGTTTCGAC 1134
QY 540 TGACTCTCATCGTATGAGCCAAAGTTTAAATCACTTTTGACAATATCTTCAGCAGATTTGAT 599
Db 1135 AGATTCAATCGTTTAACTCAACGTCGTATCCCGACAGAACAGCAGTAGAAGACTTTTAA 1194
QY 600 GGTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTGA 659
Db 1195 CATTTAATTTCCAGGAAAAAGTTTAACTGAACTTCTCTGTTTCAATTAACCAATGAAGAAGA 1254
QY 660 GACGTTGAGGTATTTTCTCACAGCAAAATCTTTTTCAGAAAGTGAACACATTTCTTT 719
Db 1255 AATGTTTGAATTCAGCATTTATGGAATCAAGTGTCTATTTTAAACAGAAACAAATGTACTT 1314
QY 720 TTATACAGCTCTTTTAGAAGGAAATTTATCCCGATACAGACCGTTTATTAATCAGCAAAAT 779
Db 1315 CTATTTCTGTTTGTAGAGGAAATTTATCTGTATACCAACCGTTTAAATTCCAACTAGCCA 1374
QY 780 TGAGACGAGGTTGTTTTCATATACCAATCCCTTCGCCACGCTATGGAACGTCGCTTCTTT 839
Db 1375 TAACAGCAAAATTTGAATTTTATGTACAGAAATTTGCTTTTCAGCAATTCGAACGTCGCTTCTTT 1434
QY 840 GATTCTTAATGCTACTCAAAATGGTACTGTTTAAAGCTTGAGATTACTCAAAATCATATTTTC 899
Db 1435 ACTTTCTCATGAAGGACGTAAACAAATATTTGTCGCTTTTCAATTTTCCACCAAGATTTCTGT 1494
QY 900 AGCTCATGTTAACTCACCTGAGTTTGGTAAGTAAACGAGGATTTAGATATTTGTTAGTCA 959
Db 1495 TTTATATGGAATTTCACTGAAATTTGGAAGTGAAGAGCTTTAAACTATGAAATGT 1554

```

RESULT 35
AL935252
LOCUS          343050 bp      DNA      linear      BCT 21-FEB-2003
DEFINITION    Lactobacillus plantarum strain WCFS1 complete genome; segment 1/11.
ACCESSION     AL935252
VERSION       AL935252.1
KEYWORDS      GI:28269805
SOURCE        Lactobacillus plantarum WCFS1
ORGANISM      Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
               Lactobacillus.
REFERENCE     1
AUTHORS      Kleerebezem,M., Boekhorst,J., van Kranenburg,R., Molenaar,D.,
               Kuipers,O.P., Leer,R., Turchini,R., Peters,S.A., Sandbrink,H.M.,
               Eijlers,M.W.E.J., Stiekema,W., Lankhorst,R.M.K., Bron,P.A.,
               Hoffer,S.M., Groot,M.N.N., Kerkhoven,R., de Vries,M., Ursing,B., de
               Vos,W.M. and Siezen,R.J.
               Complete genome sequence of Lactobacillus plantarum WCFS1
               Proc. Natl. Acad. Sci. U.S.A. 100 (4), 1990-1995 (2003)
               MEDLINE
               PUBMED
REFERENCE     2 (bases 1 to 343050)
AUTHORS      Kleerebezem,M. and Siezen,R.J.
TITLE        Direct Submission
              Submitted (14-OCT-2002) Wageningen Centre for Food Sciences, P.O.
              Box 557, 6700 AN Wageningen, The Netherlands
FEATURES     Location/Qualifiers
               1..343050
               /organism="Lactobacillus plantarum WCFS1"
               /mol type="genomic DNA"

```


10196..10798
/gene="ssb"
/locus_tag="lp_0010"
10196..10798
/gene="ssb"
/locus_tag="lp_0010"
/codon_start=1
/transl_table=11
/product="single-strand binding protein"
/protein_id="CAD62710.1"
/db_xref="GI:28269813"
/db_xref="TrEMBL:Q890K1"
/translation="MINRTILVGLTRDPELVYNGAAVATPIAVNRQFTWNGER
EADTSCIVIEKKAENFANFKHKSIVGICRITRYENQQGVRVTVTVVVENFSL
LESRAESERHOAANGSGNNYNNNGNNGNNNNNGNQGNAAPOQSSANNNNPFGN
QNTGNASSAAPSSANNNNQADPPANNQDQIDISDDLLP"
10837..11073
/gene="rpsr"
/locus_tag="lp_0011"
10837..11073
/gene="rpsr"
/locus_tag="lp_0011"
/codon_start=1
/transl_table=11
/product="ribosomal protein S18"
/protein_id="CAD62711.1"

Query Match	22.2%;	Score 252;	DB 1;	Length 343050;
Best Local Similarity	53.3%;	Pred. No. 5.1e-37;		
Matches 606;	Conservative	0;	Mismatches 520;	Indels 12; Gaps 3;
Qy	2	TGATTCAATTTTCAATTAATCGCACATTTATTTATTTTCATGCTTTAAATACAACTAAACGTG	61	
Db	1544	TTATGAAATTCACGATCAATCGATCTGCCTTCATCAAGAATTAATAACGTCCACGGG	1603	
Qy	62	CTATTAGCATAAAAATGCCAATTCCTATTTCTTTTCATCAATAAAAAATGAACTCACTTCTA	121	
Db	1604	CCATTTCTTCAAAAAACGACCAATCCAAATTTTGACTGTGCTTTAAACTTTGACGTTAATACTG	1663	
Qy	122	CAGGAGTAACTTTAAACGAGGCTTAAACGCTCAAAATATCAATTTGAAAAACACTATTTCCTGTAA	181	
Db	1664	ACGCTATTACCTTAACTGGTAGTGATGCTGNATTTTCTTTGAAAAACCACTTCAGGAA	1723	
Qy	182	GTAAAGAAATCGTGGTTTTCGTAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTAGTT	241	
Db	1724	GTGATGACAAATACGCTTAGTTCGTTGAAGATGCCGGTTCAATCGTTTTTACCTGCCGGTT	1783	
Qy	242	TTTTTATTAATATTATTTCAAGTTTGCAGATATTAGTATAAATGTTTAAAGAAATTCGAAC	301	
Db	1784	TCTTTTAGTAAATCGTAAABAAAATTACCAGAGACACCTGACTGTTTAACTGGTTCGACG	1843	
Qy	302	AACACCAAGTTGTTTAAACGAGTGGTAAATCAGAGATTACCTTTAAAGGAAAAAGATGTTG	361	
Db	1844	GTTTTTCAACACAAATTACTTCAGGGGCCGCTTCCTTTACAATCAACGGGTTAGACCCGTG	1903	
Qy	362	ACCAGTATCCTCGTCTCAAGAAGATTCACAGAAAAATCCTTTGATTTTAAAAACAATAAT	421	
Db	1904	AAAACTACCCACACTTACCTGAAATTGATACGACTAATACCACTAGCTAGCTGGGATG	1963	
Qy	422	TATTGAAGTCTATTATTGCTGAACAGCTTTTCGAGCCAGTTTCAAGAAGTCTGTCCTA	481	
Db	1964	TACTCAAGAGTTGATTGGTCAGACCGTTATTGCCGTTTCCAAATCAGGAAGTTCGACCAA	2023	
Qy	482	TTTTTAAACAGGAGTTCATATTGTTAATAGTAATCATAAAGATTTTAAAGCAGTAGTCGACTG	541	
Db	2024	TTCTGACCGGGTTTCATTTTATCTTAGCTAAACGGT--GAGTTCTTAGCGTTCGCAACCG	2080	
Qy	542	ACTCTCATCTATGAGCCAAAGTTTAACTACTTTG--GACAATACTTCAGCAGATTGGA	598	
Db	2081	ATTCAACCGCTTAAGTCAACGCCGAATCAAGTTCCAGAGGCCAATAATGCCAAATTATG	2140	
Qy	599	TGGTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCACGAGTATTTTACAGATGATTTG	658	

```

RESULT 36
AE016744
LOCUS
DEFINITION
Staphylococcus epidermidis ATCC 12228, section 1 of 9 of the
complete genome.
ACCESSION
AE016744 AE015929
VERSION
AE016744.1 GI:27314460
KEYWORDS
SOURCE
Staphylococcus epidermidis ATCC 12228
ORGANISM
Staphylococcus epidermidis ATCC 12228
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1 (bases 1 to 300787)
AUTHORS
Zhang,Y., Ren,S., Li,H., Fu,G., Lu,L., Lu,G., Jia,J., Tu,Y.,
Qin,Z., Chen,Z. and Wen,Y.
DIRECT SUBMISSION
TITLE
Submitted (05-NOV-2002) Chinese National Human Genome Center at
JOURNAL
Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES
Location/Qualifiers
1..300787
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
362..1717
/locus_tag="SE0001"
362..1717
/locus_tag="SE0001"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein"
/protein_id="AAO03598.1"
/db_xref="GI:27314461"

```

```

/translation="MGEKIDWKLVEIAQERISNTSYFTKOTQLYSLKNDIAITLV
SLPNASWLNQRYSIETQALIVDVIYGVYKCHFTSEELASVNNVNTQVEBPQVHS
SIDDTWKGEQFNHNTFTFVIGPNFRPHASLAVAEPAAEYNPLFTYGVGVGLCK
TLMHAIIGHVLNKNPKNAVITSEKKTNEFKSIKIRNETEAPREKYRILVLLIDD
IQFTQNKQEQTEBEFFTHNLFHNNKNQIVISDRPPKEIAKLEKRSRFEWGLIIVD

```


TPPYETRMALQKKEBENLDIPPEALNYIANQIOSNIRELEGATRLAYSKLOCK
PITTELAREALQIQSPKSKITIQIQKVVQQTYSVRIEDFSAKQKVENIAIPROCI
AMYLSRELTDSLKIGSEFGGRDHTTVIAHEKIANDIKSDPTFKQEVENLEKEIRN
Q"

gene

CDS

2026..3159
/locus_tag="SE0002"
2026..3159
/locus_tag="SE0002"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="AAO03599.1"
/db_xref="GI:27314462"

/translation="MMFTIKRDYFINQNTLTKAISPRITPLILTGKIDAKENEVI
LTGDSSEISITIPKQVDGEBEIVETGSLVLPGRFFVDIIKKLPKSKYKJLTNEQ
FQTILTSGHSEFNISGLDQDQYPLLEVSRRDAIQLSVKLNKNIQAQTNFAVSTSETR
PVLTCVNLIQDNELICTATDSHRLAVRLKLOLEDESENKNIIPKALSFLNKIMSDS
DEDIDIPASQNLVFRVCNINFISRLLEGHYPDTRLPPENVEIKLGINNGDFYHAID
RASLLAREGNNVILKSTGNELVELSSTSPSIGTVKEVNANDVEGGNLKISFNKYM
MDALKAIDNVEVEFFCTGMKPFILPKDDDSVTQLILPRTY"

gene

CDS

3822..4937
/locus_tag="SE0003"
3822..4937
/locus_tag="SE0003"
/codon_start=1
/transl_table=11
/product="DNA repair and genetic recombination protein"
/protein_id="AAO03600.1"
/db_xref="GI:27314463"

/translation="WKNTLOLENRYNVEQVTLDCHPVNTLIGENAGKTNLLESY
TLAKASHRTSNDKELRPKSDYAKIEGELSRYRGTMELTMFIPKKGQVKYNHLEQS
RLTQIGHNLVLFAPEDLNIVKGPSQIRRRFIDMELVQGLQISAVYINDLQAYORILKQS
NNYLKQIQGKTQTMLEVLNQFVEYALVTLRRHFIRKELETQAQTHAGITNQD
ETLIDYVPSLKSLEYANQSELIEVEIALINDLNQREKRGVCLYGPHRDDLSPNVN
GMDAQTSQOQRTALSIKLAETELMNIIEVGEYPIILLDDVLSELDDSSQTLLST
IQHKVQTPVTTTSVEGIDHEIMNNAKLRYISQGEILK"

gene

CDS

4948..6879
/locus_tag="SE0004"
4948..6879
/locus_tag="SE0004"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="AAO03601.1"
/db_xref="GI:27314464"

/translation="MVNTLSVDNNTDNYCAGQIQVLEGLAVRKRPGMYIGSTSERGL
HHLVWEIVDNSIDALAGYASHIEVIEKDNWIKVTDNGRGIIPVDIOEMGRPAVEVI
LTVLHAGKFGGKVGKSGHGVSSVWVALSODLEVYVHRNGTIVHOAYKQGVPOF
DLKEIGDRTKTGTAIRFKADKEIFTETTVNYETLQKRIELAFNLKGIQITLKDERE
SEVRDSYHGGIKSYVDLLNENKEPLHDEPIYIHOSKDDIEVEIALQYNSGYATNL
LYANNIHTYEGGTHEDGFRALTRVLNSYGTQSKIIEKDKRDSGEDTREGLTAVVS
IKHGDPOQGTQTKLGNSEVRQVVDRLFSEHFERFLYENPVSGR I IVEKIGIMASRAR
VAAKAREVTRRKSAIDVSLPGLADCCSKNPRESEIFLVEGDSAGGSTKSGRDSRT
OAILPLRGIILNVKARLDRILNNEIRQMITAGTGIGGREDISKARYHKVIMTDA
DVDGAHITLLTTFYFRMPLIEAGYVIAAPPLYKLTOQKQKYIYFNDRLEDKLKQ
ELNPSKRWISARIYKGLGEMNADQLWETTMNPEHRSMLQVRLLEDALDADQTFEMLMGDV
VENRRQFTIEDNAVYANLDF"

gene

CDS

6916..9597
/locus_tag="SE0005"
6916..9597
/locus_tag="SE0005"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit A"
/protein_id="AAO03602.1"
/db_xref="GI:27314465"

/translation="MAELPQSRINERNITSEWRSEFLDYAMSVIVSRALPDVRGLKRP
VHRRILYGLNQGMTDPKYPKQSRARIVGDVNGKYHPHGDSSIIYAMVRMAQDFSYRP
LVDDGQNGSGDGAAMRYTEARMKTITLEDRLINKDITIDFINDGNEREPFVL
PARFNNLVNAGAAGVAMATNIPPHNLTEVIDGLVLSKNPDITINELMEDIQCPDF
PTAGLVLKSGIRRAYETGRSIOWRSAEIEERGGGRQRIWVTEIPQVNAKRMIEK
IAELVRDKKIDIGTIDRLTSGTRVVIDVRKDANASVILNNLYKQTPLOTSGVYN
MIALVNGRKILNLKLEALIHYLEHQTIVRRRTNLRKADRAHILEGLQIATDHIID

EIITIRSDTDKIAMASLOERFKLTERQAQAILDMRLRLRLTGLERDKIESEYNELLE
YIKLESEIADAEVLLQVRDELTEIKERFGDERRTEIQUGLEDLEDDEDLIPREQIV
ITKLENNYIKRLPVSTRSQRRGRGIQGMNTLDEDFVSLQWTMTSHDHLFFPQNGKR
VYKLKGYEVELSRQSGKIPIINAIELENDETISTMIAVKOLESEEDYLVPATKQGI
KRSLSNFRINKNGKIAINEKEDDELIAVRLTTGNEDILIGTAHASLIRFSESLRP
LGRTAAGVKGISLREGDVTVGLDVADSESEDEVLVVTENGYGKRTPVSEYRLSNRGGK
GIKTATITERNGNIVCITTVTGEEDLMVVTNAGVIRLDVHDISQNGRAAQGVRLMKL
GDGFVSTVAKVNEEDNEENADEAQQSTTTTADVEEVDDQTPGNAIHTEGDAME
SVSPENDRIDIRQDFMVRNEDIESASDNEEDSDE"

gene

CDS

/locus_tag="SE0006"
/locus_tag="SE0006"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO03603.1"
/db_xref="GI:27314466"

/translation="MGGIYVNMKTLDSVNIPRKDETHKGDYKILLIGGSANMGCALM
LAARACVYSGSLITVATHPNHNAALHSRCPAAVIDINDTKMLTMTMEATDSILGP
GLGVDYFKGNNAITFLLQNIQPHNLIVDGDAITIFSKLQPIQPTCRVITPHQKEWER
LSGPIEBEQTYERNLEAVRIGATVILKQHGTEIYFCNDVDYKLPITGTPAMATGMGD
TLAGMITSFVQGFONFKEAVTSATYMHSEFIGEKLSDTMVYVVPSPRLISEIPVMKQLE
G"

gene

CDS

11648..12934
/locus_tag="SE0007"
11648..12934
/locus_tag="SE0007"
/codon_start=1
/transl_table=11
/product="seryl-tRNA synthetase"
/protein_id="AAO03604.1"
/db_xref="GI:27314467"

/translation="MLDIRLFRNEPEKYSIELRGDDPKVVDQVLEDEQRRELISK
TEEMKARNKYSERIAOKRKNKEDADDVIAEMRHLGDEIKDIDNOLNVDNKIRDILI
RIPNLINDEVPQASDEENVEVKWGPTRDFEPKAWHLVDELKMAADFFRAAKVSG
ARFVYLKDGALLERALNMVLTHTTQHGYTEMMTPQLVNAADMFGTQGLPKFREDL
FKVEKGLTIPTAEVPLTFNRYDEIIPQVLPFLTAQTACFACSEASAGRDTRGLI
RLHQFDKVMVRIVQPEDSDALEMTQNAEILBELGYPYRVILCTGDIQFSASKT
YDLEVLPSYNDYKEISCSNCTDQARRANIRFRDAASKPELVHTLNGSLAVGRT
FAAIVENYQEDGTLTIPEALVPPMGKTKIEKPIK"

gene

CDS

complement(13462..14466)
/locus_tag="SE0008"
complement(13462..14466)
/locus_tag="SE0008"
/codon_start=1
/product="truncated beta-hemplysin"
/protein_id="AAO03605.1"
/db_xref="GI:27314468"

/translation="MKRGVTILNQRKCILFTLLVLSLFLVFTTIVASERDFKDSL
KITTHNVYFLPTAIYPNMQSORADLISKADYIQNDVYVILNELPDKKASKRLRLH
SOYPTQPIVKGTEGQWNTSGTVRKIKVSGSGVIVSKWPIVQOEQOHIYKKGCGADM
AGNEFAYIKINKNGKYHHIIGTHLQAEADPTCFKGKOKDIRSQMSEIKQTKDNPI
KNEPVYCGDLNVIKDSDEYQOMANNLNVSLPTQPDGNAYSNDSSNSIAKYNYPKLE
PQHJDYILLDRDHAQPSWHDNTHRVKSPWVSWKGTIKYNDYSIDHYPUSGYASNE
"

gene

CDS

14829..15521
/locus_tag="SE0009"
14829..15521
/locus_tag="SE0009"
/codon_start=1
/transl_table=11
/product="amino acid permease"
/protein_id="AAO03606.1"
/db_xref="GI:27314469"

/translation="MTHLTFRQGVKECIPTLLGVAGVLSFGIIVNAAQSNFSLVLEILL
CLIIYAGAAQPIICTIATGPIAISAVITLIVNSRMFLSLTAPNTKQYGFWRKVG
LGLTLDTEFGVAITPYKGEKINDRLMHLITAYLFTWVSCVIGALFGEYISNPDA
LGLDFATIAMIEFLICISQFEGIKKSRLRIYIVLIVCVIWMMLLSSILPYSVAILIAA
IVAALLGVWVHKH"

gene

15518..15847
/locus_tag="SE0010"

[illegible]

LVDDIQFSKKEGIEEFPHFETLYNDQKQIVMTSDRLPTEIPELSERLVSFAWGL
QVEITPDPLETRIALLRKAETDGLAIDDSTLDYIASQVDYNNIRELEGALVKVQHAT
IEREDINVDLAELADLKVONRGLQSKIQEVVANYFOITSELKGRKRVQIIV
PRQIAMYSRELTDSSLPKIQEFGGKDHTTVMHACDKSRALKTDABEIKAAVYDLKA
MLEH"

gene

1546..2676

/locus_tag="LJ0002"

CDS

1546..2676

/locus_tag="LJ0002"

/codon_start=1

/transl_table=11

/product="DNA polymerase III, beta chain"

/protein_id="AA07982.1"

/db_xref="GI:41582370"

/translation="MQFTINRNLFIENLNNAMRAISSRATPIILSGIKLTITDEMLTL
TGSVDTSIEIQIDPDLLIVQSTGSIVLPARFFSEIVKLPKGKDFSEVKEFSQTKI
VSENTEFMINGLDANNYPHLEISTDAFSKISQFTREIINETVFAVATQSRPRLTG
VNFI PNNSIKAVATDSHRLSORQISLENGPOTSLDII PGKSLVELSRIIGESDPEI
TVNPGENIVFVNGNI AFYRLLDQYPTDRLIPTSTTSVEPELPVLAISLERASL
LTHERNNVKWTLDVQNLVKLQDSPEIGNVEEIEGFKNLEGDCLTISNPDPYLR
ALRASITDSIMNFTQPLRPFTVIPAKQDVNFTQLITPVRTF"

gene

2951..3100

/locus_tag="LJ0002b"

CDS

2951..3100

/locus_tag="LJ0002b"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="AA07983.1"

/db_xref="GI:41582371"

/translation="MISSGQAKWYKLENPKVKNLGEVEDRRGKKIHPGDVLNLAGEY
EFISE"

gene

3104..4228

/locus_tag="LJ0003"

CDS

3104..4228

/locus_tag="LJ0003"

/codon_start=1

/transl_table=11

/product="DNA replication and repair protein RecF"

/protein_id="AA07984.1"

/db_xref="GI:41582372"

/translation="MYLANFELKDPNFEELKINPDHVNIFIGNAQKTNLEAIY
FLALTRSHRTSDKELIRFGSKFAGLQGVKHSQVEBELKURLTPNGKKAWNRLQK
KLSAYGVGNAILFDPEDLALVKGAPISRRRPMDFEQINSEYLYFLSQVQLQQR
NNYLKQLSIKKANDLFDVLSDQLAGIAEII SRRIKYIKKLSYASAHSEISGQA
EKLOIFRPSVKETIPDDDVETIYOKVITSYKKNRPNIRKGTLSGPHRRDLDPLN
DKNADHFAQGOQRTISLSVLAELQLVHELTOEYPIILLDDVMSELDHRHQSSILNY
IHGKTQITITTDLEGISWEIVKEPKVHISAGTISTKES"

gene

4229..6196

/locus_tag="LJ0004"

CDS

4229..6196

/locus_tag="LJ0004"

/codon_start=1

/transl_table=11

/product="DNA gyrase subunit B"

/protein_id="AA07985.1"

/db_xref="GI:41582373"

/translation="WAEDKSKAALNKAQDFEKRAPTYNASOIQVLGGLLEAVKRPGMY
IGSTSSQGLHLVWEI IDNSIDERLAGFATKLEITVNEDGSGVTYQDDCGRGPVDIOAK
TGRPALETVTVLHAGGFGGGYKVGSLHGVASVNNALSTKLEVTVMRDGKKYAI
SPDHRGVGEMKQVTVPLBEHGTIVRFYPPDIFETTTTDDKILKNRIRELAPLNK
GLKLTFTDKRPSAETVYHEGGIKYEVAFNLQGEVLFDPEIYVEGNIQEDVEVA
LQYTNKYKTLMTFANNITLVEGGHEAGFKTALTRVVDYAHKAKILKDDNLSGE
DIREGITAIVSKHPNPOEGOTKTLGNSDARTAVDAFSETSRFLMENPNSVGRKI
VEKGQAEARTAAKAREVTRKKSGLAIANLPGLADNTSDPSISLFI VEGNSAG
GSAKGRSLTOAILPRLKILNVKASMDRILANQETRSUFTALGTGFGADFDVSKA
RYHKLITMDADVDAHIRTLLTLFYNMRPMIEKGXYITARPPLYQVRQCKEYL
DTDEBLGLTLPQSPRI VORYKGLGEMDPEQLMETTMNPNRRRLDRVSPYAEADA
DKIFOLLMGNEVGPREFTEKNAKYVENLDA"

gene

6207..8696

/locus_tag="LJ0005"

CDS

6207..8696

/locus_tag="LJ0005"

/codon_start=1

/transl_table=11

/product="DNA gyrase subunit A"

/protein_id="AA07986.1"

/db_xref="GI:41582374"

/translation="MANENQPDHRIIRNVDLTHMRTSFLOYAMSVIVARALPDVRDG
LKPQORRILYGMHGLVTPDKQYKKSARIIVGEVMGKFPHPGDSYILAMAHWAQDFSY
RYMLVDGHGNGFSGDGPAAWRYTEAKMSKIAVEMLRDINKETIDWQRNYDDTENEP
VVLPAFIPNLLVNGASGIAVGMTTNI PPHNLAEVIRGIMHMLMDNPVTTKOLMKVPG
PDFPTGGIVMGREGGIIRAYETKGNIVVRAKTEIETEXNGRERI VVTELPLYLVNKAEL
VKKIADLAREXTIDGITVRDESQDGMRTIIDIRRDASASVVLNNLFKETQMOANFG
MMVAVINGAPHLTLKQMLEYVLHSDVLTTRTKFELKKAEEAAHILAGLRIALDH
IDEIKIIRGSONSDIAKQILQNFGLDDRSQAILDMRLVRLTGLERDKIABEQDL
QAKIADYKDIILAKPERIDEIILNELLDIQKRDDRRRLKRIAGEAVSEDEDLISQKD
LLTLTHNGYIKRMPADKFEKQNGRGKIGKGVQDDDFINHLIFSSDHYDLFFTNL
GVLTSKAYEIPFERNAKGLPIVNLLELDKGEKIQAIVNPEGADNNLFFVTMGT
VRKTVSEFQNI RRSGLIALTRDGDLENNVLTDDGKQNLIGITHLGVAVTFENEKDV
SMGRTAGVRCGINRDNDYVVGSDILEPDSVLVISEKYGKRTAASEYPVKGRGKG
IKTANITEKNGLAGVTTVNGDEDIMLITSAGVMIRFDVDDVQSGTGRATLGVRLIKVD
DGAQVASITAVPKASDEDEESTABEEPVEN"

gene

8883..9179

/locus_tag="LJ0006"

CDS

8883..9179

/locus_tag="LJ0006"

/codon_start=1

/transl_table=11

/product="30S ribosomal protein S7"

/protein_id="AA07987.1"

/db_xref="GI:41582375"

/translation="MAKTKYEVYIIKPDIDEDSKALIDRFKVVTDNGAELESKD
WGRRFAYEIKYREGTYIHMTFVAENSEPVEFGRLSRIDNQILURSTVTKLDK"

gene

9213..9731

/locus_tag="LJ0007"

CDS

9213..9731

/locus_tag="LJ0007"

/codon_start=1

/transl_table=11

/product="single-strand binding protein"

/protein_id="AA07988.1"

/db_xref="GI:41582376"

/translation="WNNVLVGLTRDPDLRTTGGISVATFTLAVDROYTNSOGER
GADTISCVIWKAENFANFTSKSLGICIGRIQTRTYDDKDGKLVVTEVLDNFSL
LESRDRENRQTINGENFAPQGNAPSTNFGSSAFSWMNAPAGESKNKQDPFADSG
STIDISDDDLFP"

gene

9756..9992

/locus_tag="LJ0008"

CDS

9756..9992

/locus_tag="LJ0008"

/codon_start=1

/transl_table=11

/product="30S ribosomal protein S19"

/protein_id="AA07989.1"

/db_xref="GI:41582377"

/translation="MAQORRGRRRRKVDFAANHIDYDKVDVLLKRFISERGKIL
PRRVGTSAKQKRLTVAIKRVRVGLLPFFVAED"

gene

10086..11003

/locus_tag="LJ0009"

CDS

10086..11003

/locus_tag="LJ0009"

/note="Contains Mg+ and Co2+ transporter domain"

/codon_start=1

/evidence=not experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="AA07990.1"

/db_xref="GI:41582378"

/translation="MIKQKLAGNIDFYKWDYINVCDDHDIRLLKDDFDLTSEIISY
ITDHERPHFDHYITSDLLVYDVPVMTADADFTTLPIKFLMVDHTLFTVHSPDT
TYMTEEFQRDEHIHSKELIFAILFAVTFQRALSQRLVLDNNLSERHKN
DQELSOVEKSLVLSIRTNLMLESKKNKSGLHWNASEECCDDIIIEVQOSSQ
MIKYSVEITESIKSTNNLNNTNNLTQFILTWSLLLTITPTIVTGGFMVNDLPPIH
NFDWVITIVGSIIVMAILMLYLRHNNL"

[illegible]

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	66	5.8	1101	9	CNS0039G	Drosophil
C 2	63.6	5.6	603	8	BZ296086	BZ296086 CG1849.f1
C 3	63.6	5.6	614	8	BZ295031	BZ295031 CG1245.f1
C 4	57.8	5.1	1101	9	CNS000EVL	AL069706 Drosophil
C 5	57.2	5.0	1043	9	CNS0145P	AL103735 Drosophil
C 6	57	5.0	429	9	CNS003197	CG803197 1118039HO
C 7	56.2	5.0	1101	9	CNS000EVL	AL069706 Drosophil
C 8	56.2	5.0	1225	9	CNS0161D	AL106171 Drosophil
C 9	55.6	4.9	1101	9	CNS00PFG	AL071206 Drosophil
C 10	55.4	4.9	738	9	BX159596	BX159596 Danio rerio
C 11	55.4	4.9	1036	6	CD387684	CD387684 AGENCOURT
C 12	55.2	4.9	876	9	CNS009G1	AL053529 Drosophil
C 13	55	4.9	1101	9	CNS0039G	AL063921 Drosophil
C 14	54.8	4.8	1200	9	CNS016CO	AL106578 Drosophil
C 15	54.6	4.8	705	8	BH941456	BH941456 odg40C07
C 16	54.4	4.8	1061	9	CNS0015LM	AL105604 Drosophil
C 17	54	4.8	987	9	CNS0145P	AL104456 Drosophil
C 18	54	4.8	1092	9	CNS020K7	AL175696 Tetraodon
C 19	53.8	4.7	471	9	CG807194	CG807194 1118078HO
C 20	53.6	4.7	1101	9	CNS00EPA	AL069119 Drosophil
C 21	53.6	4.7	1101	9	CNS0161D	AL106896 Drosophil
C 22	53.4	4.7	1092	9	CNS020K7	AL175696 Tetraodon
C 23	53	4.7	969	9	CL469446	CL469446 SAIL 1303
C 24	52.8	4.7	1101	9	CNS001PB	AL060732 Drosophil

[illegible]

RESULT 2	BZ296086/c
LOCUS	BZ296086.c
DEFINITION	BZ296086 CG1849.f11 Candida glabrata Random Genomic Library candida glabrata genomic clone CG1849, genomic survey sequence.
ACCESSION	BZ296086
VERSION	BZ296086.1 GI:24437749
KEYWORDS	GSS.
SOURCE	Candida glabrata
ORGANISM	Candida glabrata Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida. 1 (bases 1 to 603)
REFERENCE	Wong,S., Fares,M.A., Zimmermann,W., Butler,G. and Wolfe,K.H. Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast Candida glabrata
AUTHORS	Genome Biol. 4 (2), R10 (2003)
TITLE	
JOURNAL	22508158
MEDLINE	12620120
PUBMED	Contact: Wong S Department of Genetics, Smurfit Institute Trinity College Dublin
COMMENT	

Dublin 2, Ireland
Tel: 353 1 6082319
Fax: 353 1 6798558
Email: swong@tcd.ie
Class: plasmid ends..Location/Qualifiers
1. .603
/organism="Candida glabrata"
/mol_type="genomic DNA"
/strain="CBS 138"
/db_xref="taxon:5478"
/clone="CG1849"
clone lib="Candida glabrata Random Genomic Library"

ORIGIN					
	Query Match	Score	DB 8;	Length	603;
	Best Local Similarity	49.0%;	Pred. No.	0.0012;	
	Matches	199;	Conservative	0;	Mismatches 204; Indels 3; Gaps 1;
QY	12	TTCAAATTAATCGCACATTATTATTATTCATGCTTTAAATAACAACCTAAGCGTGCCTATTAGC	71		
Dd	405	TTATATTCTCGATTCATGATATATATTTATAAAAAATTATAAAAATATTATAATATTATAATAA	346		
QY	72	TAAAAATGCCATTCCTATTCTTTTCATCAATAAAAAATTGAAGTCCACTTCTACAGGAGTAAC	131		
Dd	345	TATAAAAAATAAATAATAATATTATAGAAAAATAATTTATTAAATAATTAATAAATTATTAAC	286		
QY	132	TTTTAACAGGGTCTPAACGGGTCAAATATCAATGTGAAAAACAATTCCTGTGAAGTAATGAAAA	191		
Dd	285	TAAATCATGATTAATTTGGTTATATAGAAACTAAAAGTTCATTTTAAATTAATTTAAAAAAGA	226		
QY	192	TGCTGGTTTTGCTAAATTACCTCTCCAGGAGCTATTTTATTAGAGAGCTAGTTTTTTTATTAA	251		
Dd	225	TAATAAAATTTTCATATTNTTAAATAATTAATAATAAATAAATAGATTATTTTGTAATATA	166		
QY	252	TATTTATTCAGTTTTCCAGATATTATGATPAAATGTTTAAAGAAATTCGAACAACCACCAAGT	311		
Dd	165	TAGTATTTAAATTTATTATAAATTTAATAATTAATAATAATAATAATTTATAAAATTGAACCTATT	106		
QY	312	TGHTTTTAAACCGTGGTAAATCAGAGATTACCTTAAAGGNAAGATGTTTGACCAAGTATCC	371		
Dd	105	TGAATAATAAAGTATTGGAAT---ATATTATTAAATTTTAAAAATTTTAAATCATTAATC	49		
QY	372	TCGCTCTACAAGAGTATCAACAGAAAATCCTTTGATTTTAAAAACA	417		
Dd	48	AAGATTATTAGGTATAAAAAAGTTATGAAATCTCAATTTGAAAAAGA	3		

RESULT 3	BZ295031/c
LOCUS	614 bp DNA linear GSS 31-OCT-2002
DEFINITION	CG1245.f1 Candida glabrata Random Genomic Library Candida glabrata genomic clone CG1245, genomic survey sequence.
ACCESSION	BZ295031
VERSION	BZ295031.1 GI:24436038
KEYWORDS	GSS.
SOURCE	Candida glabrata
ORGANISM	Candida glabrata
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.
AUTHORS	1 (bases 1 to 614)
TITLE	Wong, S., Fares, M.A., Zimmermann, W., Butler, G. and Wolfe, K.H. Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast Candida glabrata
JOURNAL	Genome Biol. 4 (2), R10 (2003)
MEDLINE	22508158
PUBMED	12620120
COMMENT	Contact: Wong S Department of Genetics, Smurfit Institute Trinity College Dublin Dublin 2, Ireland Tel: 353 1 6082319 Fax: 353 1 6798558

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelOBAC11.

FEATURES

source Location/Qualifiers
 1. .1043
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone_lib="BACN11G11"
 /clone_lib="DrosBAC"
 /plasmid="pBelOBAC11"
 /note="end : T7"

ORIGIN

Query Match 5.0%; Score 57.2; DB 9; Length 1043;
 Best Local Similarity 36.2%; Pred. No. 0.035;
 Matches 145; Conservative 55; Mismatches 200; Indels 0; Gaps 0;

```

QY 29 TATTATTTCATGCTTTAATACAACTAAACGGTCTATTAGCACTAAATGCGCATTCCTA 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 TTTTCTTTTCTCTCTTTTWTATATATATATATATATATATATATATATATATATATA 594
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89 TTCTTTTCATATAAAATTAAGTGCATCTTACAGGAGTAACTTTTAAACGGGTCTAAG 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 TTTTATATATATAAAATATATATATATATATATATATATATATATATATATATATAT 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 GTCATATCAATGTAACACATATCTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 WTTTWTWTWTWTWTWTWTWTATATATATATATATATATATATATATATATATATATTA 714
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 CCTCTCCAGGAGCTATTTATTAGAGCTAGTGTATTTTATTAATATATATTTCAAGTTGC 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 TATATATTTTAAATATATATATATATATATATATATATATATATATATATATATAT 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 CAGATATTAGTATAATGTTAAAGAAATGACACACCAAGTGTGTTTAAACAGTGTGA 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 TTTTAAAAAAWAAAAAWWWAAAAAWWWAAAAAWWWAAAAAWWWAAAAAWWWAAAAA 834
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 AATCAGAGATTACCTTAAAGGAAAGATGTTGACCACTATCCTCGTCTACAAGAGTAT 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 AWATATATWAWATAWAWAWAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 894
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 389 CAACAGAAAAATCCTTTGATTTTAAAAACAAATATTGAA 428
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 895 AHAYATACAAAAACWAAATHAWAWAYAHAAWATATAA 934
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6

CG803197/c 429 bp DNA linear GSS 10-NOV-2003
 LOCUS 1118039H04.y1 1118 - RescueMu Grid S Zea mays genomic, genomic
 DEFINITION survey sequence.

ACCESSION CG803197

VERSION CG803197.1 GI:38238983

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 429)

REFERENCE

AUTHORS Walbot,V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL COMMENT

Unpublished (2001)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1118039 row: 2
 Class: transposon-tagged.
 Location/Qualifiers
 1. 429

FEATURES

source Location/Qualifiers
 1. 429
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="1118 - RescueMu Grid S"
 /note="Organ: leaf; Vector: RescueMu (engineered from
 pBlueScript backbone); Site 1: BamHI; Site 2: BglII;
 RescueMu is a 4.9 kb, modified maize Mu transposon
 designed to allow plasmid rescue from total genomic DNA.
 Mu elements insert preferentially into transcription
 units. For more information on RescueMu, go to the web
 site 'www.zmdb.laastate.edu' and follow the links for
 'RescueMu', 'Grid S' was grown at San Diego in 2002. DNA was
 extracted from leaf strips, double digested using BamHI
 and BglII, and ligated to form circular plasmids. DH10B
 cells were transformed and then screened on LB plates with
 ampicillin."

ORIGIN

Query Match 5.0%; Score 57; DB 9; Length 429;
 Best Local Similarity 49.5%; Pred. No. 0.034;
 Matches 147; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```

QY 1 ATGATTCCAATTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATACAACTAAACGT 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 ATCATATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GCTATTAGCACTAAAAATGCCATTCTCTATCTTTTCATCAATAAAAAATGAAGTCACCTCT 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 ATAATTATTATTATAATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ACAGGAGTACTTAAACAGGCTACAGGCTCAAGTCAAAATATCAATTGAAACACTATTCCTGTA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 ATAAATATAATTATAATACTATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 AGTAATGAAATGCTGTTTGTCTAATTACCTCCAGGAGCTATTATTATAGAGCTAGT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 ATTATTATAAATAATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 TTTTATTATTATTATTATTCAAGTTTCCAGATATTAGTATAATCTTAAAGAAAT 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 ATAATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 4
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7

CNS00EVL/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC:
 BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
 fly); genomic survey sequence.

AL069706

AL069706.1 GI:4949849

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

```

REFERENCE      1 (bases 1 to 1101)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL       Submitted (02-JUN-1999) Genoscope - Centre National de Sequence :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT        Web : www.genoscope.cns.fr)
               Determination of this BAC-end sequence was carried out as part of a
               collaboration with the Berkeley Drosophila Genome Project (BDGP).
               The BDGP is constructing a physical map of the Drosophila
               melanogaster genome using these BACs. For further information
               please see http://www.fruitfly.org The BDGP Drosophila
               melanogaster BAC library was prepared by Kazutoyo Osagawa and
               Aaron Mammoser in Pieter de Jong's laboratory in the Department of
               Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
               NY. The library is named RPCI-98 and was constructed by partial
               EcoRI digestion of Drosophila DNA provided by the BDGP from the
               isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
               pi and EST libraries. A more detailed description of the library
               and how to order individual BAC clones, the entire library, or
               filters for hybridization from the BACPAC Resource Center can be
               found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source           1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"

ORIGIN

Query Match          5.0%; Score 56.2; DB 9; Length 1101;
Best Local Similarity 33.9%; Pred. No. 0.059;
Matches 167; Conservative 82; Mismatches 237; Indels 7; Gaps 1;

Qy   10    TTTCCTCAATTAAATCGCACATTATTTTTTCATGCTTTTAATAACAAGCCTAAGGTCTATTAGC 69
Db                ||||| :::: |::::: |:|: |::| |::| |::| |::| |::| |::|
Qy   927   TTTTTTTTTTTWTWTTATWTTTATATWTFWWAAAAAWMAWMTWTTATTTTWTWTATW 868
Db                ||||| :::: |::::: |:|: |::| |::| |::| |::| |::| |::|
Qy   70    ACTAAAATGCCACTCTCTTCCATCAATAAAAAATGAAGTCACCTCTCACAGGAGTA 129
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::|
Qy   867   ATTWATATTAAWATTTTTTTTWTTATTTWTTATATWATWTTATWTTWTTTWAATWTTATT 808
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::|
Qy   130   ACITTAACAGGGCTAACGGTCAAATCATCAATTGAAACAACATCTCCTGTAAGAATGAA 189
Db                ||:|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|:
Qy   807   TAITWTTWTTWTTATWTTATWTTATWTTATWTTWTTWTTWTTATATATATWATATTATWTT 748
Db                ||:|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|: |::|:
Qy   190   AAATGCTGTTTGCTAAATACCCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTTTTATT 249
Db                ::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy   747   WAITWTTWTTTATWTTTNTNAWTTATWTTATWTTATWTTWTTW-----TWTTTTTTTTTT 695
Db                ::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy   250   AATATTATTTCAAGTTTGCAGATATTAGTATATAATGTTAAAGAAATTTGAACACACCAA 309
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   694   TTTTWTATTWAAATATTATTTAAATTAAWAAATWMTTATTTAATAAWTAAATTAATATTATTA 635
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   310   GTTGTTTTAACAGCTGGTAAATCAGAGATTACCTTTAAAGAAAGATGTTGACCAGTAT 369
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   634   AAAATTTTATWAAAAAATATTTTTTTTATATTAATWAWMTTATATWATTAATATATWTTTWT 575
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   370   CCTCGCTCACAAGAAGTATCAACAGAAATCCTTTTGATTTTAAAAACAAATATTATCCAAG 429
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   574   AAATTWTTAATWTTATWWWTTTAAATTTATWTTTTTTTTTAMTAAWMTWAAAWAATWAWMW 515
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   430   TCATTATTGCTGAACAGCTTTTTTGCAGCCAGTTTTCACAGAAAGCTCGCTTATTAAACA 489
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   514   WWAATAATWWWTTTWTWTTWATTWATAAATTTTWTWTTTWWAWATTWTKKKKKAADTCAGA 455
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   490   GGAGTTTCATATTG 502
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|
Qy   454   RATKTDKKKKKKG 442
Db                ||||| :::: |::::: |:~::~: ~::|: |::| |::| |::| |::| |::|

```

[illegible]

RESULT 8

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE
Drosophila melanogaster	1 (bases 1 to 1101)		Genoscope..	Submitted (02-JUN-1999)	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	Location/Qualifiers	1 -- 1101
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			Direct Submission	BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)			

FEATURES
SOURCE

ORIGIN

Query Match	4.9%	Score 55.6	DB 9	Length 1101	
Best Local Similarity	38.9%	Pred. No. 0.081			
Matches 199	Conservative	49	Mismatches 258	Indels 5	Gaps 2
QY	225	TTTATTAGAAGCTAGTTTTTTTATTAATAATATTTTCAAGTTTGCAGATATTAGTATAAA	284		
DB	1040	TATATATATWTTTATATWTTATTAATATATATATATATATTTTAMATATATWATAATAA	981		
QY	285	TGTTAAAGAARTTGAACAACACCAAGTGTGTTTAAACCAAGTGTGTAATTCAGAGATTACCTT	344		
DB	980	WATATATATTTTWTAAAAAAATTTWAATWTTTATATATATTTWTATATAATATTTATATWTTT	921		
QY	345	AAAAGGAAAAAGATGTTTGACCAAGTATCTCGCTACAGAAGTAGTCAACAGAAATCCTTT	404		
DB	920	WTATATATATAAATWTATAAAAATWTTATATATATAATWTTTTTTTAAAWATATTATTT	861		
QY	405	GATTTTAAACAATAATTTGAAGTCTATTTTGTGTAACAGCTTTTGCAGCC--AG	461		
DB	860	WATTTATATATTTWAATAATATWATATWTTTWTWTTTNTNAWTTTTTTTMMWTAATTWTWA	801		
QY	462	TTTACAGAAGTCGTCCTATTTTAAACAGAGTTCATATTGTTATTAAGTAAATCATAAAGA	521		
DB	800	TTWTATTTTATAWAWATWAAAAAATAATTTWTTTATAAAAAAWATATAA	741		
QY	522	TTTTTAAAGCAGTAGCGACTGACTCTCATCGTAGCCCAACGTTTAAATCACTTTGACACAA	581		
DB	740	WATWTATTTTATAAATAATTTAWATAAATTAATAAAAAATTAATTAATWAATATATAAA	681		
QY	582	TACTTCAGCAGATTTGATGGTAGTTCCTCCAAGTAAATCTTTGAGAGAAATTTTCAGAGT	641		
DB	680	TATATATAATTNNKKGKGGKKTBKGBBSBNNNTNTNNNTATATATATATATAAT	621		
QY	642	ATTTACAGATGATATTTCAGACCGCTGAGGTATTTTTCTCACCAAGCAATCTGTGTCAG	701		
DB	620	ATATATTTAT--TACTTAGAAATGCATCTAAATTTTATAGCCTCTAAAAATTCCTTCAA	563		
QY	702	AAGTGAACAATTTCTTTTATACAGCCTC	732		

DB	562 A A A T T A T A T T G T T A A T T G C T G T A C A G C T C 532		
RESULT 10			
EX159596			
LOCUS	EX159596	738 bp	DNA linear GSS 13-MAR-2003
DEFINITION	Danio rerio genomic clone DKEX-117F6, genomic survey sequence.		
ACCESSION	EX159596		
VERSION	EX159596.1	GI:27991215	
KEYWORDS	GSS.		
SOURCE	Danio rerio (zebrafish)		
ORGANISM	Danio rerio		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
AUTHORS	1 (bases 1 to 738)		
TITLE	Humphray S.J., Huckle.E. and Durham,J.L.		
JOURNAL	Direct Submission		
COMMENT	Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished This sequence was generated from the SP6 end of BAC 117F6. 117F6 is part of the Danokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/Projects/D_rerio/ .		
FEATURES	Location/Qualifiers		
source	1..738		
	/organism="Danio rerio"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7955"		
	/clone="DKEX-117F6"		
	/tissue_type="Testis"		
	/note="vector pindigoBAC-536"		
ORIGIN			
Query Match	4 9%	Score 55.4	DB 9: Length 738:

ORIGIN

Query Match	4.9%;	Score 55.4;	DB 9;	Length 738;
Best Local Similarity	48.9%;	Pred. No. 0.084;		
Matches 149;	Conservative 0;	Mismatches 156;	Indels 0;	Gaps 0;
Qy	49	ACAACTAAACGTCCTATTAGCAGCTAAATAATGCCATTCCCTATCTCTTTTCATCAATAAAAAATT	108	
Db	215	ACATGTCATATTGCAACAGAAATCAAGTGTTAAACATTTACTTTTGGAAATCCCTGTT	274	
Qy	109	GAAGTCACCTCTTACAGGAGTAACTTTTAAACAGGGTCTAAACGGTCAAAATATCAATTTGAAAAAC	168	
Db	275	TTTAAATATTATTAGACAAATCTTTTATCAGGAGAGAACGTATAAGAAATATTTTAAAT	334	
Qy	169	ACTATTCTCTGTAAGTAATGAAATAGCTGGTTTGCTTAATTTACCTCTCAGAGAGCTATTTTA	228	
Db	335	ACAAATCTCTTTTAAATATAAAATTTTTTTTAAAAAGATTCATATCAAAAGTAAATATTA	394	
Qy	229	TTAGAGAGCTAGTTTTTTTTTATTATATATTATTTCAAGTTTCCAGATATTAGTATAAATGTT	288	
Db	395	TTAAAAAGTTTTTAGATTTGTTTAACTATTGCTGTGTTAAATGTCTCTTTTATGTTATTTTT	454	
Qy	289	AAAGAAATTTGAACCAACCAAGTTGTTTTTAAACAGTGGTAAATCAGAGATTACCTTAAAAA	348	
Db	455	TATTATTATTATAACAAATTTTATTATATAAATGTTGTAATTCGATTTGTTTTTCCA	514	
Qy	349	GGAAA	353	
Db	515	TGAAA	519	

RESULT 11
CD387684

LOCUS	CD387684	1036 bp	mRNA	linear	EST 30-MAY-2003
DEFINITION	ACENCOURT_14346441	NIH_MGC_173	Homo sapiens	cdna 5'	mRNA sequence.
ACCESSION	CD387684				
VERSION	CD387684.1	GI:31224971			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

/note="end : T7"

```
ORIGIN
Query Match      4.8%; Score 54.8; DB 9; Length 1200;
Best Local Similarity 33.4%; Pred. No. 0.12;
Matches 115; Conservative 65; Mismatches 164; Indels 0; Gaps 0;

Qy 129 AACTTTACAGGCTTACCGTCAAAATCAATCAATGAAACCACTATCTCTGTAAGTAATGA 188
Db 836 AAGTTTTTGGGTTTWWATAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 895
Qy 189 AAATGCTGGTTGCTAAATTAATCTCTCCAGGAGCTATTTTATAGAGCTAGTTTATAT 248
Db 896 ADGAWTAATTTAAGCAAAATTTTAAWAAATTAATTTWWAAWAAWAAWAAWAAWAAWAAW 955
Qy 249 TAAATATTATTTCAAGCTTTGCCAGATATTAGTATATAATGTTAAAGAAATTTGAACAACCA 308
Db 956 TTTAATTTWTYTWTAAATTTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 1015
Qy 309 AGTTGTTTAAACCGTGTAAATCAGAGATACCTTAAAGAAAGATGTTGACCACTA 368
Db 1016 AAWTATTAAATATAAAATATAWATTTWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 1075
Qy 369 TCCTCGTCTACAGAGATATCAACAGAAATCCCTTTGATTTTAAACCAAAATATTGAA 428
Db 1076 WAAAWWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 1135
Qy 429 GTCTATTATTTGCTGAACAGCTTTTTCAGCCAGCTTTTACAAGAAA 472
Db 1136 ATTWTTTTWWAAWATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1179

RESULT 15
BH941456
LOCUS odg40c07.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.
ACCESSION BH941456
VERSION BH941456.1 GI:23421516
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 705)
Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D., and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: odg40 row: c column: 07
Seq primer: -28RPpOT reverse
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 551.
Location/Qualifiers
1. .705
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T0100DH3 buds provided by
Thomas Osborn at the University of Wisconsin.
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."
```

ORIGIN

```
Query Match      4.8%; Score 54.6; DB 8; Length 705;
Best Local Similarity 48.5%; Pred. No. 0.13;
Matches 180; Conservative 0; Mismatches 189; Indels 2; Gaps 1;

Qy 160 ATTGAAAACACTATTCTGTAAAGTAATGAATGCTGGTTTGCTTAATTAACCTCTCCAGGA 219
Db 36 ATAGCAATCAACCCACCCCTGAAGCGTCTCAAGTCTTAGTTTTTAAATTTTACTTTACAAT 95
Qy 220 GCTATTTTATTAGAAGCT--AGTTTTTTTATTAATAATTTATTTCAAGTTTGGCAGATATTA 277
Db 96 TCCGCTTTTGTGAACCTTTGAATGTTCTCAGTCTCTATATTTTATTAATTAACGGATATTC 155
Qy 278 GTATAAATGTTAAAGAAATGGAACAACCAAGTTGTTTAAACGAGTGGTAAATCAGAGA 337
Db 156 ACTTAAATGTCAAAAAACCTTAGTCACTCGAACTTTGGTTATACCGTAGTTTCTTGCAAT 215
Qy 338 TTACCTTAAAGGAAAGATGTTTCACCAAGTATCTCGTCTACAAGAAGTATCAACAGAAA 397
Db 216 GCAAGTTTAAATCATAGAGTGAATACCGCAACATACGTCGGCACCAGAACATGAGCCC 275
Qy 398 ATCCTTTGATTTTAAAAACAAAATTTATTGAAGTCTATTATTGCTGAAACAGCTTTTGAG 457
Db 276 ATATATAATCATATATATATATATAATAACATTTTTTTTTTGAAGAAATTCGCTTCAA 335
Qy 458 CCAGTTTACAAGAAGTCGCTCCTATTATTAAACAGAGTTTCATATTGTTAATAGTAAATCATA 517
Db 336 TTAACCTGCAAGAAAAAATAAAACGTTTACAAGCCCGTTGAGCTTAAAGTTAAAAATACAA 395
Qy 518 AAGATTTTAAA 528
Db 396 ACGTTTATAAA 406

RESULT 16
CNS015LM
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN14L05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL105604
VERSION AL105604.1 GI:5617618
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1061)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequences :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaut at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
Location/Qualifiers
1. .1061
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACN14L05"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

FEATURES
source
1. .1061
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACN14L05"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN
```

Query Match 4.8%; Score 54.4; DB 9; Length 1061;
 Best Local Similarity 35.6%; Pred. No. 0.15;
 Matches 143; Conservative 65; Mismatches 192; Indels 2; Gaps 1;

QY 243 TTTTATTAATATTTTCAAGTGTGCGAGATATTAGTATAAATGTTAAAGAAATTTGAACA 302
 DB 654 TTATATATATTTATTAATTTTCAAGTGTGCGAGATATTAGTATAAATGTTAAAGAAATTTMAWT 713

QY 303 ACACCAAGTGTGTTTAAACAGTGTGTTAAATCAGAGATTACCTTAAAGGAAAGAGATTTGA 362
 DB 714 AATTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAA 773

QY 363 CAGATATCTCGTCTACAGAGATTTCAACAGAAATCTTTGATTTTAAATTTTAAATTTTAAATTT 420
 DB 774 ATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 833

QY 421 TTATTTGAAGTCTATTATTTCTGTAACAGAGTGTGTTGAGAGATTTTCAAGAAATTTGCTCT 480
 DB 834 AATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 893

QY 481 ATTTTAAACAGGATTTTCAATTTTATTTAGTATAATTTTAAAGATTTTAAAGCAGTGTGCTACT 540
 DB 894 AATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 953

QY 541 GACTCTCATGATGAGCAAGTGTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 600
 DB 954 TTTTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1013

QY 601 GTAGTTCTTCCAGTAAATTTTGAAGAGATTTTCAAGCAGTA 642
 DB 1014 AATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1055

RESULT 17
 CNS014PQ/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

CNS014PQ
 Drosophila melanogaster genome survey sequence Sp6 end of BAC
 BAC1222 of Drosophila library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 AL104456
 AL104456.1 GI:5616067
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 987)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaut at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.
 Location/Qualifiers
 1..987
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BAC1222"
 /clone_lib="DrosBAC"
 /name="pBelOBAC11"
 /note="end : Sp6"

Query Match 4.8%; Score 54; DB 9; Length 987;

Best Local Similarity 32.0%; Pred. No. 0.18;
 Matches 108; Conservative 78; Mismatches 148; Indels 4; Gaps 1;

QY 152 AATATCAATTTGAAACAACTATTCCTGTAAGTAAAGAAATGCTGGTTGCTAAATACCT 211
 DB 985 AATATCACTCCMNAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAA 926

QY 212 CTCCAGGAGCTATTATTATTAAGAGTGTGTTTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTT 271
 DB 925 TTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 866

QY 272 ATATTAGTATAATTTTAAAGAAATTTGAAACAACTATTCCTGTTTAAAC-----CAGTGGT 327
 DB 865 TTWTWWAAATTTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAA 806

QY 328 AATCAGAGATTTCTTAAAGAGAAAGATGTTGACCAAGTATCTCGTCTACAAGAGTA 387
 DB 805 WAATTAATTTTAAATTTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTTAAAGAAATTT 746

QY 388 TCAACAGAAATCTTTGATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAA 447
 DB 745 WTAATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTT 686

QY 448 GCTTTGAGCAGCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTT 485
 DB 685 TTTTATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTT 648

RESULT 18
 CNS020K7/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

CNS020K7
 Tetraodon nigroviridis genome survey sequence T7 end of clone
 222L11 of library G from Tetraodon nigroviridis, genomic survey sequence.
 AL175696
 AL175696.1 GI:7813753
 GSS; genome survey sequence.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
 1
 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 20296633
 10835645
 2
 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 20359837
 10899143
 3 (bases 1 to 1092)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
 Location/Qualifiers
 1..1092

<hr/>					
ACCESSION	AL069526				
VERSION	AL069526.1	GI:4949669			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :				
	BP 191 91006 EVRY cedex - FRANCE [E-mail : seqrf@genoscope.cns.fr]				
	- Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeogawa and Aaron Mammossier in Pictet de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.				
FEATURES	Location/Qualifiers				
	1..1101				
	/organism="Drosophila melanogaster"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7227"				
	/clone="BACR29M12"				
	/clone_lib="RPCI-98"				
	/note="end : T7"				
ORIGIN					
	Query Match	4.6%;	Score 52.2; DB 9;	Length 1101;	
	Best Local Similarity	39.4%;	Pred. No. 0.46;		
	Matches 167;	Conservative 40;	Mismatches 216;	Indels 1;	Gaps 1;
QY	5 TTCAATTTTCAATTAAATCGCACATTAATTTATTTCATGCTTTTAAATACAACATAACGTGGCTA	64			
Db					
	TTTTWATTTTATATATTTTWTWATTTTAWATTTATTTAAATTTWWTTTWMTTWTATTTT	1028			
QY	65 TTAGCACTAAAAATGCCATTCCTATTCCTTCATCATATAAAAATGAAGTCACCTCTACAG	124			
Db					
	TTTTTWTNNAATWMAWTAWTTTWWTTTAAATTTTAAATTTTATATATAATWAAATTT	968			
QY	125 GAGTAACCTTTAACAGGGTGCTACGGGTCAAATTCAAATTCGAAACACACTATTCCCTGTAAGTA	184			
Db					
	WTWTTTNTTAAATATAATTTTATAAATAATTTATTTWATNTNAMAATWATTTAWAAAAAA	908			
QY	185 ATGAAAATGCTGGTTTGCTAAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTT	244			
Db					
	AAAAAAAATWTTTWTATTTTWTAAAAATATA-AAAAATAATTTATTATAAAATTTTTTTT	849			
QY	245 TTATTAAATATTTATTCGAAGTTGGCCAGATATTAGTATAAATGTTAAAGAAATTTGAACAAC	304			
Db					
	TTTTTWTTTTAAATTTATATAAAAAAATAWTTTWTAWATAAAWAAAAATTTTAAATTTATCATA	789			
QY	305 ACCAAGTGTGTTTAAACCAGTGTTGAATTCAGAGATTACCTTTAAAGGAAGAAGATGTTGACC	364			
Db					
	MTTWATWTTWTAATAAAATTTTWTWTATTTTAAAAATTTTWTWAAAAAATAAAAGM	729			
QY	365 AGTATCCTCGTCTACAGAAGATTCACAGAAATCCTTTGATTTTAAAAACAAAAATTTAT	424			
Db					
	CWTCMTATATWAAAAAATAAAAAAAAAATAWAAAAAATTTNTYTITTTTAAWAAAAA	669			
QY	425 TGAA 428				

QY 155 TATCAATTGAAACACATATTCCTGTAAGTAATGAAAAATGCTGGTTTGCTAATTACCTC 214

ORIGIN	Query Match	4.5%;	Score 51.4;	DB 8;	Length 575;
	Best Local Similarity	49.1%;	Pred. No. 0.64;		
	Matches 136;	Conservative 0;	Mismatches 141;	Indels 0;	Gaps 0;
Qy	1	ATGATTCAATTTTCAATTAATCGCACATTATTTATTTCATGCTTTAAATACAACTAAACGT	60		
Db	277	ATTATTATTATTAGTATTATTATTAGTATTATTATTAGTATTATTATTACTACTATTATT	218		
Qy	61	GCTATTAGCACATAAAATGCCATTCCTATTCTTTTCATCAATAAAATGGAAGTCACCTCT	120		
Db	217	ACTATTATTATTATTAGTATTATTATTAGTATTATTATTATTATTACTATTATTACTACT	158		
Qy	121	ACAGGAGTAGTACTTTAACAGGGTCTAACGGTCAAAATATCAATTGAAAACACTATTTCCTGTA	180		
Db	157	CCATTATTACTATTAGTATCATTTATTATTATTATTGCTATTATTACTATTATTATT	98		
Qy	181	AGTAATGAAAATGCTGGTTTGCTPAAATACCTCCAGGAGCTATTTTATTAGAAGCTAGT	240		
Db	97	AGTATTATTATTACTATTATTACTACTACTACTACTACTAGTATTATTATTATTATTATT	38		
Qy	241	TTTTTTTATTAAATTAATTCAAGTTGGCCAGATATTA	277		
Db	37	ATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT	1		

RESULT 31	CNS0124K	1085 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	CNS0124K				
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN07F22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL101102				
VERSION	AL101102.1	GI:5612713			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster	(fruit fly)			
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1085)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant . The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pSheloBAC11.				
FEATURES	Location/Qualifiers				
source	1..1085				

```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN07F22"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : T7"

ORIGIN

Query Match          4.5%;   Score 51.4;   DB 9;   Length 1085;
Best Local Similarity 24.9%;   Pred. No. 0.7;
Matches 125;   Conservative 115;   Mismatches 263;   Indels 0;   Gaps 0;

Qy      132  TTTTAAACAGGTCCTAACGGTCAATATCAATTGAAACACTATTTCCTGTAAGTAATGAAA 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      523  KBNBNNNNKNNNAANNAANKKCAAAAAAANNAANNAANNAANNAANNAANNAANNAAN 582

```

192	Qy	TGCTGGTTTGCTAAATTACCTCTCCAGGAGCTATTATTATAGAAGCTAGTTTTTTTATTAA	251
583	Db	NAKNNNNAANVNAANNNNNAAAAAATBAAATTATTAAGAAATATATTTATTTT	642
252	Qy	TATTATTCCAAAGTTTGCAGATATTTAGTATATAAATGTTAAAGAAATTTGAACAACACCAAGT	311
643	Db	TWTAATATATANKTTATAAKTTTAKATKAKAKKKTCKAKAKAKAKTAKTAAKADKAA	702
312	Qy	TGTTTTTAACCAGTGGTATAATCAGAGATTTACCTTAAAAAGGAAAAAGATGTTGACCAAGTATCC	371
703	Db	KDKTKTKKKKKKKKAGKKKAAKTAKRTAKAGAAAKAGKAGAKAKKRAAAAKAT	762
372	Qy	TCGTCTCAAGAAGTATCAACAGAAAAATCCTTTGATTTTAAAAACAAAAATTTATGAAGTC	431
763	Db	KAAKTAKATAAKKAATKTAAAKKAADATKTTTNTTKTDIAKTTTAAAKATTTTKTATTK	822
432	Qy	TATTATTGCTGAAACAGAGCTTTTGCAGCAGTTTTACAGGAAGTCGTCCTTATTTTAAACAGG	491
823	Db	TTDKAAAKKWAAKAAGWAKAKKAKKKKTGTAGARAKAKAGAKAAKAKKKKKKK	882
492	Qy	AGTTCATATTGTATTAGTATATCATAAAGATTTTAAAGCAGTAGGCACTGACTCTCATCG	551
883	Db	AAAKAGKKKKKDKAKKKKTAAAKKTTKTKKKKTAKTKAAGGGGAAAGAGGGGAGKGR	942
552	Qy	TATGAGCCAAACGTTTTAATCTACTTTTGACAATATCTTCAGCAGATTTTGATGGTAGTCTCTCC	611
943	Db	AAAGTGAKRKKKDKATAAAGWAAAAAAATKKTKADWATADKKTKKKTAKAAKT	1002
612	Qy	AAGTAAATCTTTGAGAGATTTT	634
1003	Db	KAAWAKAAKTATAAADADDDK	1025

RESULT 32
CNS00FYG
LOCUS
DEFINITION
CNS00FYG 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR32N04 of RPCI-98 library from *Drosophila melanogaster* (fruit
fly), genomic survey sequence.

AL071206	
AL071206.1	GI:4951245
GSS	
Drosophila melanogaster (fruit fly)	
Drosophila melanogaster	
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
(bases 1 to 1101)	

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 131 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial, EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

FEATURES
source
1..1101
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

/db_xref="taxon:7227"
/clone="BACR32N04"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
    Query Match
    Best Local Similarity 4.5%; Score 51.2; DB 9; Length 1101;
    Matches 157; Conservative 47; Mismatches 235; Indels 0; Gaps 0

QY 90 TCTTTTCATCAATATAAAATTGAAGTCACCTCTACAGGAGTAACCTTTAACAGGCTCTAACGG 149
Db 618 TATATTATATATATATATANNNNNANNNNWNNVVVVVMMCCMCCMNNNAATTATAT 677

QY 150 TCAATATCAATTGAAAAACACTATTCTGTGAAGTAATGAAATGCTGCTTTGCTTAATTAC 209
Db 678 ATATTTTATATATTWATTAAATTATTTTAAATTATTTATTTATTTTATATAAAATAW 737

QY 210 CTCCTCCAGGAGCTATTTTATTAGAGCTAGTTTTTTTATTATAATTATTTTCAAGCTTTGCC 269
Db 738 ATWTTATATWTTTTTTTATAANAWAATATTTTTTTTTTTTTTTTWWTAWTTWTATATAATA 797

QY 270 AGATATTAGTATAAATGTTTAAAGAAATTGACAACACCAAGTGTGTTTTAACCACTGGTAA 329
Db 798 WAAATWAAATATAWWAAAAAATAWNNAAAAMWAAWATATWAAATTTTAAATATATAWA 857

QY 330 ATCAGAGATTACCTTTAAAGGAAAGATGTTTGACCAAGTATCCTCTCAACAAGAAAGTATC 389
Db 858 ATWAAATAATATWTTTTAAAAAAWATTWTATATATAWAAWTTTTTATWAAATTTATATA 917

QY 390 AACAGAAAACTTTGATTTTAAAAACAAAATTATTGGAAGCTATTTATGCTGGAACAGC 449
Db 918 TAWAAAWTATAWATAATTATATAWAWATATATAAAWATTAWAATTTTTTAWAAATATAT 977

QY 450 TTTTGCAGCCAGTTTACAAGAAAGTCCTCTATTTTAAACAGGAGTTCATATTGTATTAAAG 509
Db 978 ATWTTTATATWTTATATWTAATAATATATAWATATATATTATTAATAWAAAAWATATAT 1037

QY 510 TAATCATAAAGATTTTAAA 528
Db 1038 ATATATATAWAAATATAATA 1056

```

RESULT 33	CG807323/c				
LOCUS	CG807323	400 bp	DNA	linear	GSS 10-NOV-2003
DEFINITION	1118080G01.y1 1118 - RescueMu Grid S Zea mays genomic survey sequence.				

ACCESSION CG807323
VERSION CG807323.1 GI:38244959
KEYWORDS GSS.

SOURCE	ORGANISM
Zea mays	Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 400)
Wallbot. V.

TITLE
 JOURNAL
 COMMENT
 Maize genomic sequences found using engineered RescueMu transposon
 Unpublished (2001)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1118080 column: 3
 Class: transposon-tagged.

```

FEATURES
source
  transposon-tagged.
  Location/Qualifiers
    1..400
    /organism="Zea mays"
    /mol_type="genomic DNA"

```


/cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="1118 - RescueMu Grid S"
 /note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 4.5%; Score 50.6; DB 9; Length 400;
 Best Local Similarity 48.4%; Pred. No. 0.91;
 Matches 140; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCATTAATCGCACATTTATTATTCATGCTTTAAATACAACTAAACGT 60
 DB 331 ATTATTATTATTAATATTATTATTAATTTATTATTAATTTATTATTAATTTATTAAT 272
 QY 61 GCTATTAGCAGTAAATGCGCATTCCTATTCTTCATCAATAAAATGAAAGTCACTTCT 120
 DB 271 ATAATATTATTATTAATTTATTATTAATTTATTATTAATTAATTAATTTATTATTAAT 212
 QY 121 ACAGAGTAACCTTTAACAGGGTCTAACGGTCAAAATATCAATTTGAAAACACTTTCCTGTA 180
 DB 211 ATAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 152
 QY 181 AGTAATGAATGCTGGTTGCTAATACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 DB 151 ATAATTAATACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 92
 QY 241 TTTTATTATTAATTAATTTCAAGTTGCGCAGATATTAGTATATAAATGTTA 289
 DB 91 ATAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43

RESULT 34

CG803196
 LOCUS 1118039H04.1EL_x1 1118 - RescueMu Grid S Zea mays genomic, genomic survey sequence.
 DEFINITION

ACCESSION CG803196
 VERSION CG803196.1 GI:38238982

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 430)

Walbot,V.

Maize genomic sequences found using engineered RescueMu transposon

Unpublished (2001)

CONTACT Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Very probable ligation site of ends cut by single endonuclease.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1118039 row: 2

Class: transposon-tagged.

FEATURES

source

Location/Qualifiers

1..430
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="1118 - RescueMu Grid S"
 /note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 4.5%; Score 50.6; DB 9; Length 430;
 Best Local Similarity 48.1%; Pred. No. 0.92;
 Matches 143; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 1 ATGATTCAATTTTCATTAATCGCACATTTATTATTCATGCTTTAAATACAACTAAACGT 60
 DB 130 ATTATTATTATTAATTTATTATTAATTTATTATTAATTTATTATTAATTTATTATTAAT 189
 QY 61 GCTATTAGCAGTAAATGCGCATTCCTATTCTTCATCAATAAAATGAAAGTCACTTCT 120
 DB 190 ATTATTATTATTAATTTATTATTAATTTATTATTAATTTATTATTAATTTATTATTAAT 249
 QY 121 ACAGAGTAACCTTTAACAGGGTCTAACGGTCAAAATATCAATTTGAAAACACTTTCCTGTA 180
 DB 250 ATAATTTATTATTAATACTATCCCTATAATTAATAAATAATTTGTTTCTTATCCCTT 309
 QY 181 AGTAATGAATGCTGGTTGCTAATACCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 DB 310 ATTGTTAAAAAAAATTTATTATTTATTATTAATTAATTAATTTATTATTAATTTATTAT 369
 QY 241 TTTTATTATTAATTTCAAGTTGCGCAGATATTAGTATATAAATGTTAAAGAAATTT 297
 DB 370 ATTATTATTATTTATTATTAATTTATTATTTATTATTTATTATTTATTATTTATTATTTAT 426

RESULT 35

BJ414423

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BJ414423 528 bp mRNA linear EST 10-MAR-2002
 BJ414423 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddv19i06 5', mRNA sequence.

BJ414423

EST.

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

```

source
1. 528
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv19106"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

```

Query Match	4.5%;	Score 50.6;	DB 4;	Length 528;	
Best Local Similarity	46.6%;	Pred. No. 0.95;			
Matches 158;	Conservative	0;	Mismatches 181;	Indels	0; Gaps 0;

QY	23	GCACATATTATTTCATCGCTTTAAATACACTAAACGTGCTATTAGCACTAAAAATGCCA	82
Db	2	GCATATAATTNATTCAATACATTTAATNAAACAATGAATTCCTACTCGTAATCACTTTTCAT	61
QY	83	TTCTCTATTCTTTTCATCAATAAAAAATTCGAAGTCACCTTCACAGAGTAACCTTTTAACAGGGT	142
Db	62	TAGTTCGTCGTTATTCAACTTCAGAAGGTGTHAAATCGCTTCCATGCTTCCACATCAA	121
QY	143	CTACGGTCAAAATATCAATTGAAAAACAATTCCTGTAAGTAATGAAAAATGCTGTTTGC	202
Db	122	CTGCTTTAAAGATTGAAAAATGAAAAATTTACAGATAAAATTTCAATTAAGAATGTTTATG	181
QY	203	TAAATTACCTCTCCAGGAGCTATTTTATTAGAGCTAGTTTTTTTAAATTAATATTATTTCAA	262
Db	182	TGGAATGTTTACATGCACCAATTAATCCACGAGATTTAAATATTATTTCACGGTACTTATG	241
QY	263	GTTTGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACAACACACCAACGAGTGTGTTTAACCA	322
Db	242	GTACAAATGTTTCAAGTTGGAGGTGTAGCAGGTATGGAAGGTGTTGGTGTGTTTAAAAAAGG	301
QY	323	GTGCTAAATCAGAGATTACCTTAAAAAGGAAAAAGATGTTG	361
Db	302	TTGCTAGTGGTGTTCACAGAGTTTAAAGAGAAATGATTAG	340

RESULT 36
CNCN0145P/c
LOCUS
DEFINITION
 CNCN0145P 1043 bp DNA linear GSS 26-JUL-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC
 BACN1G11 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL103735
 AL103735.1 GI:5615346
 GSS.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1043)
REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaut at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBel0BAC11.

```

source
location/qualifiers
1..1043
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

```

```

/clone="BACN11G11"
/clone lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN

Query Match      4.5%  Score 50.6;  DB 9;  Length 1043;
Best Local Similarity 35.3%;  Pred. No. 1;
Matches 85;  Conservative 49;  Mismatches 107;  Indels 0;  Gaps 0;

QY      198  TTTCGCTAAATTACCTCTCCAGGAGCTATTTTATTAGAAGCTAGTTTTTTTATTAAATATTAT 257
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      854  WTWTWTATTTWATATATWTARRGGRADDAAAAAAAATTTTWTWWWWTTTWWW 795

QY      258  TTCAAGTTTGCCAGATATTAGTATTAATGTTTAAAGAAATTGAACAACACCAAGTGTGTTTT 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      794  WWWWWTTTTTTTTTTTAAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 735

QY      318  AACCACTGGTAAATCAGAGATTACCTTAAAGCGAAAGAGATGTGACCAAGTATCTCCGCT 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      734  TATTAATTTWAAAWATAWATATAATATATWATAAAAAAATTTTATWAWAWWWATAATATA 675

QY      378  ACAAGAAGTATCAACAGAAAAATCCTTTGATTTTAAAAAACAATATTATTGAAGTCTATTAT 437
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      674  WAWAAAAAWAWAWAWAAAAAWATATATTTTWWAWATAAAWAWAWAAAAAATAWATAWA 615

QY      438  T 438
      |
Db      614  T 614

```

RESULT 37	
LOCUS	CG745316/c
DEFINITION	1193 bp DNA linear GSS 24-OCT-2003
ACCESSION	P038-1-G09.ya Ppa EcorI BAC Library Pristionchus pacificus genomic,
VERSION	genomic survey sequence.
KEYWORDS	CG745316
SOURCE	CG745316.1 GI:37966242
ORGANISM	GSS.
REFERENCE	Pristionchus pacificus
AUTHORS	Pristionchus pacificus
TITLE	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus. 1 (bases 1 to 1193)
JOURNAL	Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J.
MEDLINE	An integrated physical and genetic map of the nematode Pristionchus pacificus
PUBLISHED	Mol. Genet. Genomics 269 (5), 715-722 (2003)
COMMENT	128835951 12884007 Contact: Sommer RJ

```

evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
Class: BAC ends.
      Location/Qualifiers
        1..1193
          /organism="Pristionchus pacificus"
          /mol_type="genomic DNA"
          /strain="California"
          /db_xref="taxon:54126"
          /clone_lib="Ppa EcORI BAC Library"
          /note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
FEATURES
source

```

Query Match 4.5%; Score 50.6; DB 9; Length 1193;
Best Local Similarity 45.7%; Pred. No. 1.1;

19	AATCGCATTATTATTTCATGCTTTAAATCAACTAAACGTGCTATTAGCACTAAAAAT	78
17	AACAGCATATAATTCATTTCATACATTTAATCAACACATGAATTTCTACTCGTAATATCATTT	76
79	GCCATTCCATTCCTTCATCAATAAAAATTGAACTCACTTCTACAGGAGTAACCTTTAAACA	138
77	TCATTAGTTCGTGCTGTTATTCAACTTCAAGAAAGTGTAAATCGCTTCACATGGTTCCACA	136
139	GGGTCTAACGGTCAAAATATCAATTCGAAACACTATTCTCGTAAGCTAATGAAAAATCGCTGGT	198
137	TCAACTGCTTTAAAGATTGAANAATGAAANAATTTACAGATAAAATTTCAAATAAAGATGTT	196
199	TTGCTAATTAACCTCTCCAGGAGCTATTTTATTTAGAAGCTAGTTTTTTTATTTAAATATTATT	258

Db 197 TTAGTTGAAATGTTTATCATGCGCAATTAATCCAGCAGATTTAAATATATTATTCAAGTACT 256
 Qy 259 TCAAGTTTCCGCGATATTAGTATAAATGTTAAAGAAATCGAACACACCAAGTTGTTT 318
 Db 257 TATGTCACAAATGTTCAAGTTGGAGGTGTAGCAGGTATGGAAGTGTGGTGTGTTTAA 316
 Qy 319 ACCAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTTG 361
 Db 317 AAGTTGGTAGTGGTGTGTACAGGATTTAAAGAGAAATGATTAG 359

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2005, 11:56:08 ; Search time 494 Seconds
(without alignments)
13189.954 Million cell updates/sec

Title: US-10-048-071-27
Perfect score: 1134
Sequence: 1 atgtattcaatttcaataa.....ttacaccagtcgacaaac 1134

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

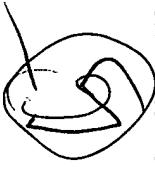
Result No.	Score	Query Match	Length	ID	Description
1	1126	99.3	1137	16	US-10-282-122A-38117
2	690.8	60.9	1134	16	US-10-282-122A-35834
3	608.2	53.6	21338	8	US-08-961-527-20
4	603.4	53.2	1137	16	US-10-158-844-20
5	603.4	53.2	1137	9	US-09-815-242-9348
6	603.4	53.2	1137	16	US-10-282-122A-37498
7	341.2	30.1	1131	9	US-09-815-242-6810
8	338	29.8	1128	16	US-10-282-122A-20800
9	330.6	23.2	8001	9	US-09-070-927A-622
10	315.6	27.8	1128	16	US-10-282-122A-21823
11	247.4	21.8	1131	16	US-10-282-122A-34649
12	243.4	21.5	1131	16	US-10-282-122A-35724

13	236.2	20.8	1134	9	US-09-815-242-8628
14	236.2	20.8	1134	17	US-10-449-830A-1
15	236.2	20.8	1134	8	US-08-781-986A-153
16	236.2	20.8	2347	16	US-10-329-624-153
17	233	20.5	1134	9	US-09-815-242-8602
18	233	20.5	1134	15	US-10-282-287-9
19	233	20.5	1134	16	US-10-282-122A-7900
20	230	20.3	1128	9	US-09-815-242-4614
21	200.4	17.7	4736	16	US-10-398-221-3763
22	198.8	17.5	684707	16	US-10-398-221-9
23	198.8	17.5	3011208	16	US-10-398-221-2058
24	185.6	16.4	1146	16	US-10-282-122A-24149
25	178.4	15.7	1137	16	US-10-282-122A-9444
26	105.8	9.3	1072	9	US-09-974-300-1554
27	103.2	9.1	1206	16	US-10-282-122A-17164
28	102.4	9.0	1125	16	US-10-282-122A-16193
29	88	7.8	992	16	US-10-671-403-173
30	88	7.8	992	16	US-10-671-419-173
31	88	7.8	992	16	US-10-670-844-173
32	88	7.8	992	16	US-10-671-134-173
33	88	7.8	992	16	US-10-673-098-173
34	88	7.8	992	16	US-10-672-638-173
35	88	7.8	992	16	US-10-673-127-173
36	88	7.8	992	17	US-10-670-817-173
37	88	7.8	992	17	US-10-673-119-173
38	88	7.8	992	18	US-10-671-207-173
39	74	6.5	1101	16	US-10-282-122A-15836
40	68.6	6.0	486	9	US-09-974-300-6041
41	64.2	5.7	482	16	US-10-282-122A-4503
42	64.2	5.7	482	16	US-10-282-122A-4503
43	59.2	5.2	640681	9	US-09-790-988-1
44	58.4	5.1	1193	16	US-10-282-122A-9797
45	56.2	5.0	1501	18	US-10-473-126-328

ALIGNMENTS

RESULT 1

US-10-282-122A-38117
; Sequence 38117, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27



; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38117
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-38117

Query Match 99.3%; Score 1126; DB 16; Length 1137;
Best Local Similarity 99.6%; Pred. No. 8.8e-225;
Matches 1129; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATACAACTAAACGT 60
DB 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATACAACTAAACGT 60
QY 61 GCTATTAGCACTAAATGCGATTCCTATTTCTTTCATCAATAAATTCGAAGTCATCTT 120
DB 61 GCTATTAGCACTAAATGCGATTCCTATTTCTTTCATCAATAAATTCGAAGTCATCTT 120
QY 121 ACAGGAGTAACTTTAACAGGGTCTAAGCGTCAAAATATCAATTTGAAACACTATTCCTGTA 180
DB 121 ACAGGAGTAACTTTAACAGGGTCTAAGCGTCAAAATATCAATTTGAAACACTATTCCTGTA 180
QY 181 AGTAATGAAATGCTGGTTGCTAATTAATCTCCAGGAGCTATTTTATTAAGCTAGT 240
DB 181 AGTAATGAAATGCTGGTTGCTAATTAATCTCCAGGAGCTATTTTATTAAGCTAGT 240
QY 241 TTTTATTATTAATTAATTTCAAGTTTGCAGATATAGTATAAATCTTTGATTTTAAACAAAA 300
DB 241 TTTTATTATTAATTAATTTCAAGTTTGCAGATATAGTATAAATCTTTGATTTTAAACAAAA 300
QY 301 CAACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTAAAGGAAAAAGATGTT 360
DB 301 CAACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTAAAGGAAAAAGATGTT 360
QY 361 GACCAATATCTCTGCTCAACAGAGTATCAACAGAAATCTTTGATTTTAAACAAAA 420
DB 361 GACCAATATCTCTGCTCAACAGAGTATCAACAGAAATCTTTGATTTTAAACAAAA 420
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCGCT 480
DB 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTCAAGAAAGTCGCT 480
QY 481 ATTTTAAACAGGAGTTTCAATTTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
DB 481 ATTTTAAACAGGAGTTTCAATTTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT 540
QY 541 GACTCTCATCGTATGAGCAAGCTTTAATCACTTTGGACAATCTTCAAGAGATTTTCAATG 600
DB 541 GACTCTCATCGTATGAGCAAGCTTTAATCACTTTGGACAATCTTCAAGAGATTTTCAATG 600
QY 601 GTAGTTCTTCCAGTAATCTTTGAGAAATTTTCAAGAGATTTTCAAGATGATTTAG 660
DB 601 GTAGTTCTTCCAGTAATCTTTGAGAAATTTTCAAGAGATTTTCAAGATGATTTAG 660
QY 661 ACCGTGAGGATTTTCTCCACCAAGCAATCTTTGTCAGAGTGAACATTTCTTTT 720
DB 661 ACCGTGAGGATTTTCTCCACCAAGCAATCTTTGTCAGAGTGAACATTTCTTTT 720
QY 721 TATACACGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGATTT 780
DB 721 TATACACGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGATTT 780
QY 781 GAGCGGAGGTTGTTTTCATACCAATCTCTTCCGACCGTATGAAAGCTGCTCTTCTG 840

DB 781 GAGCGGAGGTTGTTTTCATACCAATCTCTTCCGACCGTATGAAAGCTGCTCTTCTG 840
QY 841 ATTTCTTAATGCTACTCAAAATGCTTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
DB 841 ATTTCTTAATGCTACTCAAAATGCTTACTGTTAAGCTTTGAGATTTACTCAAAATCATATTTCA 900
QY 901 GCTCATGTTAACTCACTGAGGTTGTTAAGGTTAAACGAGGATTTAGATATTGTTAGTCAG 960
DB 901 GCTCATGTTAACTCACTGAGGTTGTTAAGGTTAAACGAGGATTTAGATATTGTTAGTCAG 960
QY 961 TCTGCTAGTGTATTTAACTATCAGCTTCAATCAACTTACCTTATTTAGATCTTTTAAAGCT 1020
DB 961 TCTGCTAGTGTATTTAACTATCAGCTTCAATCAACTTACCTTATTTAGATCTTTTAAAGCT 1020
QY 1021 ATTTAAAGTGAACACAGTAAATTAATTTCTTATCACCAGTTCCGACCATTCACCTTAACA 1080
DB 1021 ATTTAAAGTGAACACAGTAAATTAATTTCTTATCACCAGTTCCGACCATTCACCTTAACA 1080
QY 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1134
DB 1081 CCAGCGATGAGGAGAAAGTTTATCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1134

RESULT 2

US-10-282-122A-35834
; Sequence 35834, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35834
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Streptococcus mutans
US-10-282-122A-35834

Query Match	60.9%;	Score	690.8;	DB	16;	Length	1134;
Best Local Similarity	75.6%;	Pred. No.	5e-134;				
Matches	857;	Conservative	0;	Mismatches	277;	Indels	0;
Gaps	0;						
Qy	1	ATGATTCAATTTTCAATTAAATCGCACATTAATTTATTCATGCTTTAAATACAACCTAAACGT	60				
Db	1	ATGATATAATTTTCAATTAATAAAGTTTTTTTTTTTCAAGCCTTAATGCTACCAACGGA	60				
Qy	61	GCTATTAGCACTAAAAATGCAATTCCTATTCTTTTCATCAATAAAAAATGGAAGTCACCTTCT	120				
Db	61	GCTATTAGTCTAAAAATGCTATTCTCTATTCTTTCTAGTTTAAAAATGGAAGTGAATTTCT	120				
Qy	121	ACAGGAGTAACCTTTAACAGGGCTTAACGGGTCAAAATATCAATTTGAAAAACAATAATTCCTGTGA	180				
Db	121	CAGTCCATTACTTTTAAACAGGCTCTAATGAGCAAAATTTCTATTGAAAAATACGATTTCAAGCT	180				
Qy	181	AGTAATAGAAAATGCTGGTTTGCTAAATTTACCTCTCCAGGAGCTATTTTATTATAGAGCTAGT	240				
Db	181	GAAGAAGAAATGCTGGACTATTGGTTACTTCTTCAGGAGCTATCTTACTTTGAAGCTAAT	240				
Qy	241	TTTTTTTATTAATATTATTTTCAAGGTTTGCCAGATATTAGTATAATGTTAAAGAAATTTGAA	300				
Db	241	TTCTTTATTAATATTGTTTCAAGCCTACTGATATCACTTTAGATTTTGAAGAAATTTGAA	300				
Qy	301	CAACACCAAGTTGTTTTTAAACAGTGGTAAATACAGAGATTACCTTTAAAGGAAAGATGTT	360				
Db	301	CAACATCAAGTTGTGTTTTTAAATAGTGGAAATCAGAAATTAACCTTTAAAGGTAAAGATGTG	360				
Qy	361	GACCAGTATCCTCGTCTACAAGAGTATCAACAGAAATCCTTTGATTTTAAAAACAAAA	420				
Db	361	GAACAAATATCCTCGTTTGCAGGAAGTTGGAACAAATTAATCCTTTAAATTTTAGAAACAAAA	420				
Qy	421	TTATTGAAGTCTATTATTGCTGAAAACAGCTTTTTCAGCCAGTTTTCACAGAAAGTCGTCCCT	480				
Db	421	TTGTTAAAAACAATATTCTGAAAAACGCTTTTGTCTAGTACCCAAAGAAAGTCGACCA	480				
Qy	481	ATTTTTAACAGAGTTTCATAATGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT	540				
Db	481	ATTTTAAACAGGTGTTCACTTTGGTATTGACAAATCATAAAGAGTTTAAAGCTGTGCTACA	540				
Qy	541	GACTCTCATCGTATGAGCCAAAGTTTTAAATCACITTTGACAAATACTTCAGCAGATTTCATG	600				
Db	541	GACTCTCATCGCATGAGTCAAGAAAAATTAACCTTCGACCATTTCATCAGATGATTTTGAT	600				
Qy	601	GTAGTTCTTCCAAAGTAAATCTTTTGAGAGAAATTTTTCAGCAGTATTTCACAGATGATATTGAG	660				
Db	601	GTGCTCATTCGAAGCCGTTCTTTACGTGAATTTGCGAGCTGATATTACAGATGATATTGAA	660				
Qy	661	ACCGTTGAGGTATTTTTCTCACCAAGCCAAATCTTTGTCAGAAAGTGAACATTTCTTTT	720				
Db	661	TCGTGTGAGGTATTTTTCTCTAACAGTCAAAATCCTCTTTAGAAGTGAATATATTAGCTTC	720				
Qy	721	TATACAGCCTCTTAGAAGGAAATTAATCCCGATACAGACCGTTTATTAAATGACAGAAATTT	780				
Db	721	TATACGCGCTGTTTAGAAGGAAATTAATCTTGATACTGATCGCTGTTTGGGCAATAACTTT	780				
Qy	781	GAGACGAGGTGTGTTTTCAATACCAATCCTTCGCCACGCTATGGAACGTCCTCTCTG	840				
Db	781	GAACAAGAGTGTGTTTCAATACATATATGCTCTCTCGTTATGCTATGGAGCGTCTCATTTA	840				
Qy	841	ATTTCTAATGCTACTCAAAATGGTACTGTTTAAGCTTGAGATTAATCAAAATCATATTTTCA	900				
Db	841	ATCTCAAAATGCTACCCAAATGGTACTGTTTAAATTTGGAGATTAATCAACAATCAAGTAACA	900				
Qy	901	GCTCATGTTTAACTCACTCGAGGTTGGTAAAGTAAACGAGGATTTAGATATATTGTTAGTCAG	960				
Db	901	GCACATGTTAAATTCACCTGAGGTTGGAAAAAGTCAATGAAGAACTTGATATTGGAAGCTTA	960				
Qy	961	TCTGGTAGTGAATTTAACTATCAGCTTCAATCCCACTTACCTTATTGAGTCTTTTAAAAAGCT	1020				
Db	961	TCAGGAAATGATTTTAAACAAATAGCTTTTAAACCACTTACTTGATCGAAGCTTTTAAAAAGCT	1020				
Qy	1021	ATTTAAAGTGAACAGTAAAAAATTCATTTCTTATCACAGTTTCAGCAATTCACCTTAACA	1080				

[illegible]

Qy	301	CAACACCAAGTGTGTTTTTAACCAAGTGGTAAATCAGAGATTACCTTTAAAGAGAAAAGATGTT	360
Db	7376	CAAAATCAAATGTGTTTTTAACCAAGTGGCAAAATCAGAAATTAACCTTAAAGGAAAGATAGC	7435
Qy	361	GACAGTATCTCGTCTACAGAAGTATCAACAGAAATCTTTTGAATTTTAAAAACAAA	420
Db	7436	GAAACAATCCAGGAATCCAAGNAATTTTCAGCAAGCACTCTTTAAATCTTGAACAAA	7495
Qy	421	TTATTGAAGTCTATTATTGCTGAAACAGCTTTTCAGCCAGCTTTTACAGAAAGTCGTCCT	480
Db	7496	TTACTCAAGAAAATATTAAATGAACACGCTTTGCTGCAAGTACACAAGAGAGTCGTCGG	7555
Qy	481	ATTTTACAGGAGTTCATATTGTATTAAAGTAATCATAAAGATTTTAAAGCAGTAGCGACT	540
Db	7556	ATTTTACAGGTGTCACCTTCGTATTGAGTCAACACAAAGAGTTTAAAAAACAGTTTGCAACA	7615
Qy	541	GACTCTCATCGTATGAGCCACGTTTAAATCACTTTGGACAATACTTCAGCAGATTTGATG	600
Db	7616	GACTCTCATCGCTTAAGCCAGAAAAAATGACTCTTGAAAAAATAGTGATGATTTTGTAT	7675
Qy	601	GTAGTCTTCCAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTCACAGATGATATTGAG	660
Db	7676	GTCTGTAATTCCTAGCCGTCTCTACGCGAAATTTTCAGGGTATTTCACAGATGATTCGAA	7735
Qy	661	ACCGTTGAGGTATTTTCTCACCAAGCCAAATCTTGTTCAGAGTGAACAATTTCTTTT	720
Db	7736	ACTGTAGAGATTTTCTTTGTCNAATTAACCAATCTCTTTAGAAGCGAAAATATTAGCTTC	7795
Qy	721	TATACAGCCCTCTTAGAAGGAAATTAATCCCGATACAGACCGTTTATTTAATGACAGAAATTT	780
Db	7796	TATACTCGTCTCTAGAAGGAATCTCTCTGATACAGATCGTTTGAATCCAAACAGACTTT	7855
Qy	781	GAGACGGAGTGTGTTTTCAATACCCAAATCCCTTCGCCAGCTATGGAACGTGCTTCTTG	840
Db	7856	AACTACTATTACTTTTAAATGTGTTAAACTTACGCCAGTCAATGGAGCGTGCCCGTCIT	7915
Qy	841	ATTTCTAATCGTACTCAAAATGGTACTGTGTTAAAGCTTGAGATTACTCAAAATCATATTCA	900
Db	7916	TTATCAAGTCGACTCAAAATGGTACTGTGAAACTTTGAAATTAAGATGGGGTGTGTAGC	7975
Qy	901	GCTCATGTTAACTCACTGAGTGTGGTAAAGGTAAACGAGGATTTAGATATTGTTAGTCAG	960
Db	7976	GCCCATGTTCACTCTCCAGAAGTTGGTAAAGTAAACGAAGAAATCGATCTGATCAGTT	8035
Qy	961	TCGTGTAGTGAATTAATCATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTAAAGCT	1020
Db	8036	ACTGGTGAATTTGACCAATTAGTTTCAACCCCAACTTACTTGATTGATTTCTCTTAAAGCT	8095
Qy	1021	ATTTAAAGTGAACAGTAAAAATTTCAATTTCTATCACCAGTTTCGCACATTTCAACCTTAACA	1080
Db	8096	TTAAATAGCGAAAAGGTGACTATTAGCTTTATCTCAGCTGTTGCTCCATTTACTCTGTG	8155
Qy	1081	CCAGGCGATGAGGAAGAAAGTTTTTATCCAAATTAATTTACACCAAGTACGAACAA	1133
Db	8156	CCAGCAGATCTACGAAGAAGCTTCATGCGAGCTCATTTACACCAAGTTCGTACAA	8208

RESULT 4

```

US-10-158-844-20
; Sequence 20, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
;
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```


7676	GT	CGTAATTCCTAGCCGTTCTCTACGGAAATTTTCACGGGTATTTACAGATGATATCGAA	77335
661	QY	ACCGTTGAGGATATTTTTCTCACCAAGCCAAATCTTGTTCAAGAAGTGAACACATTTCTTTTT	720
7736	Db	ACTGTAGAGATTTCTTTGGCAATAACCAAAATCCTCTTTAGAAGCGAAAATATTAGCTTC	7795
721	QY	TATACAGCCCTCTTAGAAGGAATATCCCGATACAGACCGTTTATTATATGACAGAAATTT	780
7796	Db	TATACTCGTCTCTAGAAGGAATCTATCTTGATACAGATCGCTTGATTCGAACAGACTTT	7855
781	QY	GAGACGGAGGTGTGTTTTCAATACCCAAATCCCTTTGCCACGCTATGGAACGTGCTTTCTTG	840
7856	Db	AACACTACTATTTACTTTTAAATGTGGTAAACTTAGCCAGTCAATGGAGCGTCCCGTCTT	7915
841	QY	ATTTCTTAATGCTACTCAAAATGCTACTGTTAAAGTTGAGATTACTCAAAATCATATTTCA	900
7916	Db	TTATCAAGTCGCACTCAAAATGGTACTGTGAAACTTTGAAATTAAGGATGGGGTTGTAGC	7975
901	QY	GCTCATGTTAACTCACCTGAGGTTGGTAAAGTAAAACGAGGATTTAGATATGTTAGTTCAG	960
7976	Db	GCCCATGTTTCACTCTCCAGAAGTTGTTAAAGTAAAACGAAGAAATCGATCTGATCAGGTT	8035
961	QY	TCGTGTAGTGATTTAACTATACGTTTCAATTCGAACTTACCTTATATGAGTCTTTAAAGCT	1020
8036	Db	ACTGGTGAAGATTTGACCAATAGTTTCAACCCCAACTTACTTGAATGATTTCTTCTTAAGCT	8095
1021	QY	ATTTAAAGTCAAAACAGTAAAAATTCATTTCTTATACACAGTTCGACCATTACCCCTAAACA	1080
8096	Db	TTAAATAGCGAAAGGTGACTATTAGCTTTATCTCAGCTGTTTGTCCAATTTACTCTTG	8155
1081	QY	CCAGGCGATGAGGAAGAAAGTTTTATCCAAATTAATACACAGTACGAACAAA	1133
8156	Db	CCAGCAGATACTGACGAAGACTTCATCGAGCTCAATACACAGTTCGTAACAA	8208

RESULT 5

```

US-09-815-242-9348
; Sequence 9348, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/131,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9348
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae

```


; Sequence 6810, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlseen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6810
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1131)
US-09-815-242-6810

Query Match 30.1%; Score 341.2; DB 9; Length 1131;

Best Local Similarity 57.0%; Pred. No. 4e-61;
Matches 644; Conservative 0; Mismatches 483; Indels 3; Gaps 1;

```
Oy 4 ATTCAATTTTCAATTAATGCGACATTATTATTATGCTTTTAAATACAACTAAAGTGCT 63
Db 1 ATGAATTTACGGTCAAGGAAGTCTTTTACGAAGATTACAACCTGTTCAACGACA 60

Oy 64 ATTAGCACTAAAAATGCCATTCCTATTCTTCATCAATAAAAAATGAAAGTCATCTTACA 123
Db 61 ATTTCTTCTAAAAACAACGATTCCAATCTTAACAGGTGTTAAAAATGTTGTATCAGAAGAT 120

Oy 124 CGAGTAACCTTTAACAGGCTCAACGGTCAATATCAATTAATGAAACACATATTCCTGTAAGT 183
Db 121 GCGTTATCACTTACTGTTGGAGTAAACCGGATATTTCAATTTGAAAGTTTTTAAAGTAAAGAT 180

Oy 184 AATGAAATGCTGGTTGCTTAATACCTCTCCAGCAGCTATTATTTATAGAAGCTAGTTTT 243
Db 181 GATGAAAGCTCAATGACCATTTGAACGCACAGTAGCATCGTTTTACAATCTCGTTTC 240

Oy 244 TTTATTAATATTATTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAATTTGAACAA 303
Db 241 TTGCGGAAATTTATTCGTTAAATTTACCAAGATATGTTTCAATGGAAGTTCTAGATAAC 300

Oy 304 CACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTTGAC 363
Db 301 AATCAAGTTGCAATTTACTTTCAGGAAAGCTGATTTTACGGTTAATGGAATGATGCTGAC 360

Oy 364 CAGTATCTCGTCTACAAGAGTATCAACAGAAAAATCCTTTGATTTTAAAAACAAATTA 423
Db 361 AATTATCGCACTTACCAAGTAATGATCTCATAAACCAAAATGAAATTTACCTGTTCAATTA 420
```

```
Oy 424 TTGAAGTCTATTATTGCTGAAACAGCTTTTTCAGCCAGCTTTTCAAGAAAGTCGTCCTATT 483
Db 421 TTGACAAAAATATTAGTGAACAGGTTTTCCTGTATCGATGCACGAAAGTCGTCCTAATT 480

Oy 484 TTAACAGAGTTTCATATTGTTAATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGCTGAC 543
Db 481 TTAACAGAGTTTCATATTGTTAATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGCTGAC 537

Oy 544 TCTCATCGTATGAGCCAAACGTTTAAATCACTTTTGGCAATATCTTTCAGCAGATTTTGTGTA 603
Db 538 TCACATCGTTTAAAGTCAACGTTGTATCCGACAGAAACAGCAGTAGAAGACTTTAAACATT 597

Oy 604 GTTCTTCAAGTAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATTTGAGACC 663
Db 598 GTAATTCAGGAAAAAGTTTAACTGAACTTTCTCGTTTCAATTAACCAATGAAGAAAGATG 657

Oy 664 GTTGAAGTATTTTCTCACCAGGCAAACTCTGTTTTCAGAGTGAACACATCTTTTAT 723
Db 658 GTTGAAGTATTTTCTCACCAGGCAAACTCTGTTTTCAGAGTGAACACATCTTTTAT 717

Oy 724 ACAGCGCTCTTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTTGAG 783
Db 718 TCTGTTTGTAGAAGGAAATTTATCCTGATACCAACCGTTTAAATTCCAACTAGCCATAAC 777

Oy 784 ACGAGGTTGTTTTCATATACCAATCCCAATCCCTTCGCCACGCTATGGAACGTCCTTTT 843
Db 778 ACGCAATTTGAATTTTATGTACCAAGATTTGTTTTCAGCAATCGAACGTCCTTTTACT 837

Oy 844 TCTAATGCTACTCAAAATGTTGTTTAAAGCTTGAAGTAAACGAGGATTTAGATATTGTTAGT 903
Db 838 TCTCATGAAGACGTAAATATTTGTTGCGCTTTCAATTTCCAGATTTCTGTTGTTTAA 897

Oy 904 CATGTTAACTCACCTGAGGTTGGTAAAGTAAACGAGGATTTAGATATTGTTAGTCACT 963
Db 898 TATGGAATTTCACTGAAATTTGGAAGAGTTCGAAGAGCTTTTAACTATGAAAAATGTTTCT 957

Oy 964 GGTAGTGAATTAATCACTATCAGCTTCAATCCAACTTACCTTATGATCTTTTAAAGCTATT 1023
Db 958 GGTGAAGCGTTGGATATTTCTTTCAACCCAGATATATATGAAGATGCGTTGCGAGCGTTT 1017

Oy 1024 AAAAGTGAACAGTAAAAATTTCTTTTATCACCAGTTTCAGCCATTTCACCTTACACCA 1083
Db 1018 GGCATATGAATATATCCGTGAATTTCTTTTCAATTTCTGTTTACATTTGAGGCA 1077

Oy 1084 GGCATATGAAGTAAAGTTTATCCAAATTAATTAACCCAGTACGAAACAAA 1133
Db 1078 ACCGAACAGAACTAGATTTCATTTCACTAATTTACACCGGTTTCGTACAAA 1127
```

RESULT 8

US-10-282-122A-20800

; Sequence 20800, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlseen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20800
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-282-122A-20800

Query Match          29.8%; Score 338; DB 16; Length 1128;
Best Local Similarity 56.8%; Pred. No. 1.8e-60;
Matches 642; Conservative 0; Mismatches 485; Indels 3; Gaps 1;

QY 4 ATTCAATTTTCAATTAATCGCATATTTATTTATTCATGCTTTTAAATACAACTAAAGTGT 63
DB 1 ATGAAATTAACGGTCAACGAAGTGTCTTTTACGAAGATTACAACTGTTTCAACGAGCA 60

QY 64 ATTAGCACTAAAATGCCATTCCTATTTCTTCAATCAATAAAAAATGAAAGTCACTTCTACA 123
DB 61 ATTCTCTTAAACAACGATTCCAACTTAACAGGTGTAAATTTGTGTATCAGAAGAT 120

QY 124 GCGATTAATTTAACAGGCTCAACGGTCAATATCAATTAATGAAACACATTTCTGTAAAT 183
DB 121 GCGTTATCACTTACTGTTGGGAGTAAACGGGATATTTCAATTTGAAAGTTTAAAGTAAAGAT 180

QY 184 AATGAAATGCTGGTTTGTCTAATTAACCTCCAGAGCTATTTTATTAAGAGCTAGTTT 243
DB 181 GATGAAAGCTCAATGACCATTAACGACGACGATAGTCTGTTTCAATCTCGTTTC 240

QY 244 TTTATTAATATTTTCAAGTTTGCAGATATTTAGTATAAATGTTTAAAGAAATTTGAA 303
DB 241 TTTGCGGAAATTTATTCGTAATTTACGAAAGATATGTTTCAATGGAAGTTCTAGATAAC 300

QY 304 CACCAAGTTGTTTAAACAGTGTAAATCAACAGATTTACCTTAAAGGAAAGATGTTGAC 363
DB 301 AATCAAGTTGCAATTTACTTCAAGAAAGCTGATTTTACGGTTAATGGAATGATGCTGAT 360

QY 364 CAGTATCTCGTCTACAGAAAGTATCAACAGAAATCTTTGATTTTAAACAAATAA 423
DB 361 AATTATCGGCACTTACCAAGTATTTGATATCTCAAAACCAAAAGAAATTTACCTGTTCA 420

QY 424 TTGAAGTCTATTTGCTGAACAGCTTTTCAGCGAGTTTACAGAAAGTGTGCTTAT 483
DB 421 TTGCAAAATTTATTTAGTGAACAGGTTTCTGCTGATGATGATGATGATGATGATGAT 480

QY 484 TTAACAGAGTTCATATTTGATTAAGTAAATCATATAAGATTTTAAAGCAGTAGGACGAC 543
DB 481 TTAACAGGTTTCACTTTATTTTGAAGAAATCAAAA--TTACTTGGCTTGGCAGAGAT 537

QY 544 TCTCATGATGACGCAACGTTTAACTCACTTTGACAAATATCTCAGCAGATTTGATGTA 603
DB 538 TCACATCGTTTAAAGTCAACGTTGATCCGACAGAAACAGCAGTAGAAGACTTTAAACAT 597

QY 604 GTTCTTCAAGTAAATCTTTGAGAGATTTTTCAGCAGTATTTTACAGATGATTTGAGACC 663
DB 598 GTAATTCAGGAAAGTTTAACTGTTTCTGTTTCAATTAACCAATGAAGAAAGT 657

```

```

QY 664 GTTAGAGTATTTTCTCACCAGCAAAATCTGTTTCAAGAGTGAACACATTTCTTTTAT 723
DB 658 GTTGAATCAGCATTTATGGAATAATCAAGTGCTATTTAAACAGAAACAATGTACTTCTAT 717
QY 724 ACACGCTCTTTAGAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTTGAG 783
DB 718 TCTCGTTTGTAGAGGAAATTTATCCTGATACCAACCGTTTAAATTTCCAACTAGCCATAAC 777
QY 784 ACGGAGGTTGTTTCAATPACCAATCCCTTCGCCACGCTATGGAACGTCCTTCTTTGATT 843
DB 778 ACGCAATTTGAATTTTATGTACCAGAAATTTCTTCCAGCAATCGAACGTCCTTCTTACTT 837
QY 844 TCTAATGCTACTCAAAATGGTACTGTTTAAGCTTGAGATTTACTCAAAATCATATTTTACGCT 903
DB 838 TCTCATGAAGGACGTAACAATATTTGTCGCTTTCAATTTCCACGAGATTTCTGTTGTTTAA 897
QY 904 CATGTTAACTCACCTGAGGTTGTAAGTAACAGAGATTTAGATATTTGTTAGTCAGTCT 963
DB 898 TATGMAATTCACCTGAAATTTGGAAGAGTGAAGAGCTTTAACTATGAATAATGTTTCT 957
QY 964 GGTAGTGATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGAGCTATT 1023
DB 958 GGTGAAGCGTTGATATTTCTTTCACCCAGATTTATATGAAGATGCGTTGCGAGCGTTT 1017
QY 1024 AAAAGTGAACAGTAAATAATTTCTTCTTATCACCAGTTGACCAATTCACCTTAACACCA 1083
DB 1018 GGCATATGATATTTCCGTTGAATAATTTCTTCTCAATTCGTCGTTTACATTTGAGGCCA 1077
QY 1084 GGCATGAGGAAGAAAGTTTATCCAAATTAATTAACACAGTACGAAACAAA 1133
DB 1078 ACCGAAACAGAACTAGATTTCAATTTCACTAACTAATTAACACCGGTTCTGTACAAA 1127

```

RESULT 9

```

; Sequence 622, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
;             Patrick J. Dillon
;             Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

```

TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 622:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8001 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 622:

US-09-070-927A-622

Query Match 29.2%; Score 330.6; DB 9; Length 8001;
Best Local Similarity 56.9%; Pred. No. 1.5e-58;
Matches 645; Conservative 0; Mismatches 485; Indels 4; Gaps 2;

QY 1 ATGATTCAATTTTCAATTAATCGCACATTAATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB ATTATGAAATTAACGGTCNAACGAAGTGTCTTTTACAAGAATTACAACTGTTTCAACGA 657

QY 61 GCTATTAGCACTAAAATGCCATTCCTATTCTTT-CATCAATAAAAAATGAAGTCACCTTC 119
DB GCAATTTCTTCTAAACCAACGATTCCTCAATCTTAACAGGTGTATAAATTTGTTTATCAGA 717

QY 120 TACAGGAGTAACCTTTAACAGGGTCTAACGGTCAATATCAATTTGAAACACACTATTCCGT 179
DB AGATGGCTTATCACTTACTGAGGATACCGGATATTTCAATTTGAAGTTTTTTAAGTAA 777

QY 180 AAGTAATGAAATGCTGGTTTGTCTAATTAACCTCTCCAGGAGCTATTTTATTAGAGCTAG 239
DB AGATGATGAAAAGCTCAAAATGACCATTTGAACGCACAGGTAGCATCGTTTACAACTCG 837

QY 240 TTTTATTATTAATTTATTTCAAGTTTGCAGATATTAGTATAATGTTAAAGAAATTTGA 299
DB TTTCTTTTGGCGAAATTTATTCGTAATTTACCAAGATATGTTTCAAAATGGAAGTTCTAGA 897

QY 300 ACAACACCAAGTTGTTTAAACAGTGGTAAATTCAGAGATTACCTTAAAGGAAAGAAAGT 359
DB TAAACATCAAGTTGCAATTAATCTCAGGAAAGCTGATTTACGGTTATGATAGATAGC 957

QY 360 TGACAGATATCTCGTCTACAAGAAGTATCAACAGAAAATCTTTGATTTTAAACACAAA 419
DB TGATAATTTATCGCACTTACCAGTAATGATATCAAAACCAAAATGAAATTTACCTGTTCA 1017

QY 420 ATTATTGAAGTCTATTATTTGCTGAAACAGCTTTTCCAGCCAGTTTACAAGAAAGTCGCC 479
DB TTTATTGCAAAAATTTATTTAGTGAACACAGGTTTTCGTATCGATGCAAGAAAGTCGCC 1077

QY 480 TATTTTAAACAGGATTCATTTATTTATTAAGTAAATCAATAAGATTTTAAAGCAGTAGCGAC 539
DB AATTTTAACTGGGTTCACTTTATTTTAGAAAATCAAAA---TTACTTGGCGTTGCGAC 1134

QY 540 TGAATCTCATCGTATGAGCCAACTTTTAACTTATTCATCTTTGGACATATCTTCAGCAGATTGAT 599
DB AGATTACATCGTTTAAAGTCAACGGTGTATCCGACAGAAACAGCAGTAGAAGACTTTAA 1194

QY 600 GGTAGTTCTTCAAGTAATCTTTGAGAGAAATTTTCCAGCAGTATTTACAGATGATTTGA 659
DB CATTTGTAATTCAGGAAAAGTTTAACTGAACCTTTCTCGTTCAITTAACCAATGAAGAAGA 1254

QY 660 GACCGTTCAGGTATTTTCTCACCAAGCCAAATCTTTGTCAGAAGTGAACACATTTCTTT 719
DB AATGTTTGAATTCAGCATTTATGGAATAATCAAGTGTATTTTAAACAGAAACAAATGACTT 1314

QY 720 TTATACAGCCTCTTTAGAGGAAATTTATCCCGATACAGACCGTTTATTAATGACAGAAAT 779
DB CTATTTCTGTTTGTAGAGGAAATTTATCTGATACCAACCGTTTAAATTCGAATGACCA 1374

QY 780 TGAGACGGAGGTTGTTTTCATACCAATCCCTTCGCCACGCTATTTGGAACGTCCTTT 839
DB TTAACACGCAAAATGAAATTTTATGTATCCAGAAATTTGCTTTTCCAGCAATTCGAACGTCCTTT 1434

QY 840 GATTTCTTAATGCTACTCAAAATGGTACTGTTTAAAGCTTGAGATTACTCAAAATCATATTTC 899
DB ACTTCTCATGAAGACGCTAACAAATATTGTTCCGCTTTTCAATTTTCCACAGATCTGTTGT 1494

RESULT 10

US-10-282-122A-21823

; Sequence 21823, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 21823

; LENGTH: 1128

; TYPE: DNA

; ORGANISM: Enterococcus faecium

US-10-282-122A-21823

Query Match 27.8%; Score 315.6; DB 16; Length 1128;

Best Local Similarity 55.6%; Pred. No. 8.7e-56;

Matches 628; Conservative 0; Mismatches 499; Indels 3; Gaps 1;

QY 4 ATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGTGCT 63

Db 1 ATGAAAGTACTTTTAAACCGAGCTAGCTTTATGAGGAATGCAAACTGTTCAACAGCT 60
 QY 64 ATTAGCACTAAATGCAATCTCTATTTCTTTATCAATATAAATGGAAGTCACTTCTACA 123
 Db 61 ATTTCAAGCAAAACCGAGTCCCTATTTTGACAGGTGTAAAAATCACTGACACAAGAA 120
 QY 124 GGAAGTAACTTTAAACAGGCTTAACGGTCAAAATATCAATGAAACACTTCTCTGAAGT 183
 Db 121 GGTGTTGACTTTGACGGGAGCAAGCTGATATATCAATGAAACTTTTGTCTGTGAA 180
 QY 184 AATGAAATGCTGGTTGCTAAATTAACCTCTCCAGGAGCTATTTTATTAGAAGCTAGTTTT 243
 Db 181 AACGAAAGCAAAATATGCAATCGAATCTACTGGTCCCATTTTTCACAGCAGTTTC 240
 QY 244 TTTTAAATATATTTCAAGTTGCGAGTATCACTGAGGAGTATTTTACTTTAGAAATTTAGAAAAT 300
 Db 241 TTTAGGCAATCAATCGGAGACTTCTGAGGAAACATTTTACTTTAGAAATTTTAGAAAAT 300
 QY 304 CACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGATGTTGAC 363
 Db 301 AAACAAGTAGCGATCACTCTCGAAAGCGAATTTTATCGTAATGATAGATGAGAT 360
 QY 364 CAGTATCTCTGCTACAGAGTATCAACAGAAATCTCTTGTATTTTAAAAACAAATTA 423
 Db 361 AACTATCTCTCATCTCTGTTGCGAAAGCCATAACAGATGAAATTTACCTGTACAGTA 420
 QY 424 TTGAAGTCTATTTCTGCAACAGCTTTTGCAGCCAGTTTACAGAAAGTCTCTTAT 483
 Db 421 TTGACTAACTAATCAACGAAACAGTTTGTGCTCTCAACATGAGAGTCTGCTCAATC 480
 QY 484 TTAACAGGAGTCTATTTGATTTATTAAGTAAATCAATAAGATTTTAAAGCAGTAGCAGTCA 543
 Db 481 TTGACAGGTGTCATTTCAATTTA---TCTGATAATCTTTTATTAGCTGAGTACTGAT 537
 QY 544 TCTCATGCTATGAGCAAGTCTTATCACTTTGGCAATATCTTACAGAGATTTGAGGTA 603
 Db 538 TCTACCGTCTAAGTCAACGCGTGAATCCAGTAGAACAAGCGGCTGATCATTTTGTATTT 597
 QY 604 GTTCTTCCAGTAAATCTTTGAGAGAAATTTTACAGAGATTTTACAGATGATATTGAGACC 663
 Db 598 GTTATTTCTGGAAAAAGTTTGTAGTCAATTTCTCGCTATTAACAATGAGAGAAATC 657
 QY 664 GTTGAGTATTTTCTCACCAGCAAAATCTTGTGTTGAGAGTGAACACATTTCTTTTAT 723
 Db 659 GTCGAAATCAGCATTTATGAAACCAAGTGTGTTTCAAAACAGAAACGATGATTTCTAT 717
 QY 724 ACACGCTCTTGAAGGAAATTTATCCGATACAGACCGTTTATTAATGACAGAAATTTGAG 783
 Db 718 TCCCGTTTCTGAGAGGAAATCTCCAGATACCAATCGTTTGTGATTTCTTCAAGTTTAA 777
 QY 784 ACGGAGTGTGTTTCAATACCAATCTCTTCCGACGCTATGGAACGTCCTTCTTGTAT 843
 Db 778 ACAGAGTGTGAATTTCTGTTCTAGCTTTTATAGCGGATCGAAGCGCTCTTTTACTT 837
 QY 844 TCTATGCTACTCAAAAGTGTACTGTTAAGCTTGAGTCTGAGTACTCAAAATCATATTTAGCT 903
 Db 838 TCTATGAGCGCGTGAACAATCGTCTGTTTATCTATTTCTGTCAGATGCGGTTGTTT 897
 QY 904 CATGTTAACTCACTGAGTGTGTAAGGTAACAGGATTTAGATATTTGTTAGTCACTCT 963
 Db 898 TACGAAATTCGCCAGAAATCGGAAAGTAGAAGAGCTTGAATTTATACAGCAAGCAGT 957
 QY 964 GGTAGTGAATTAATCACTAGCTTCAATCCAACTTACCTTATTTAGTCTTTTAAAGCTATT 1023
 Db 958 GCGCATCAATTAGATATTTCTTCAATCCGATTTATATGAAGCAGCGTTGCGTGCATTT 1017
 QY 1024 AAAAGTGAACAGTAAATTAATTTCTTTATCAACAGTTCGACCATTCACCTTAACACCA 1083
 Db 1018 GGAGATGAGCATCAAGTGAATTTATCTCTGCTATTTGCTCATTTTATGAGTGAACCA 1077
 QY 1084 GCGGATGAGGAGAAAGTTTATCCAAATTAATTAACAGCAGTACGAAACAA 1133

Db 1078 ACGGAAGATGGCTTCAATTTATTTCAGCTGATTACTCTGTGACGACAAA 1127

RESULT 11
 US-10-282-122A-34649
 ; Sequence 34649, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 34649
 ; LENGTH: 1131
 ; TYPE: DNA
 ; ORGANISM: Staphylococcus epidermidis
 ; US-10-282-122A-34649

Query Match 21.8%; Score 247.4; DB 16; Length 1131;
 Best Local Similarity 52.9%; Pred. No. 1.4e-41;
 Matches 580; Conservative 0; Mismatches 511; Indels 6; Gaps 2;

QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATCAACTAAACGT 60
 Db 1 ATGATGGAATTTCAATTAAGAGATTTATTTATTAATCAACTTAACGACATTTAAAG 60
 QY 61 GCTATTAGCACTTAAATGCCATTTCTTATTTTTCATCAATAAAATTTGAAGTCACTTCT 120
 Db 61 GCTATCTCCCTAGAACAACTTTTACCGATTTTAAACAGGTATCAAAATTTGATGCTAAAGAA 120
 QY 121 ACAGGAGTAACTTTAACAGGCTTACGGTCAAAATTAATCAATTAATCAACTATTCCTCC 177
 Db 121 AATGAAGTCAATTTTAACTGGTTCAGATTCAGAAATATCAATAGAAATCACTATTCCTCAA 180
 QY 178 GTAAGTAAATGAAATCTGTTTGTCTTAAATTTACTCTCCAGGAGCTATTTTATTAGAGCT 237
 Db 181 CAAGTTGATGCTGAGAAATTTGTAATTAACAGAAACAGGATCACTTCTCTCTCTGCT 240

```
QY 238 AGTTTTTTTAAATATTTTCAAGTTTCCAGATATATAGTATAAATGTTAAAGAAAT 297
Db 241 CGTTTCTCGTTGATATTTAAATAAACTTCTCGAAAGAAGTTAAATATCAACTAAT 300
QY 298 GAACAACACCAAGTTGTTTAAACCAAGTGTAAATCAGAGATTACCTTAAAGGAAAGAT 357
Db 301 GAACAATTCACCAACGTTAATACATCAGTCAGTCATTCAGATTTAACTTAAGTGGCTTAGAT 360
QY 358 GTTGACCAAGTATCTCGTCTACAAGAAGTATCAACAGAAATCTCTTTGATTTTAAACA 417
Db 361 CCTGATCAATACCCATTTATCTCAGGATATCAGGATGAGCTTATCAATTTGTCAGTT 420
QY 418 AAATTTTGAAGTCTATTTATCTCAACAGCTTTTTCAGCAGCTTTTACAGAAAGTGT 477
Db 421 AAGTACTTAAATAATATCATCTGCAACAATTTTTCAGTGTCCACTCAGAAACACGA 480
QY 478 CCTATTTTAAACAGGAGTTTATATCTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTTACTGGTGAACCTGGCTTATA---CAAGATATGAATTAATATGCACAGA 537
QY 538 ACTGACTCTCATCTGATGAGCAACGTTTATCACTTTGGACAATPACTTTCAGCAGATTTG 597
Db 538 ACAGATTCACACCGCTTAGCTGTAAGAAAGTTTACAGTTAGAGATGAATCAGAAATATA 597
QY 598 ATGTTAGTTCTTCAAGTAATCTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATATT 657
Db 598 AATGTCATCATCTCTGGTAAGCTTTTATCTGAAATTTAAACAAATTTATGAGTGACAGCAC 657
QY 658 GAGACCGTTGAGGTTTCTTCTCACAAGCCAAATCTTTGTCAGAAAGTGAACACATTTCT 717
Db 658 GAAGATATTGATTTCTTCTGCTTCTAACCAAGTGTATTTCAGAGTGGGNAATTAAT 717
QY 718 TTTTATACACGCTCTTGAAGAAATATATCCGATACAGACCGTTTATTAATGACAGAA 777
Db 718 TTTATCTCACGTTTACTTGAAGTCAATATCCAGATACGACACGCTTTATCCAGAAAT 777
QY 778 TTTGAGCGGAGTTGTTTTCATACCCAAATCCCTTCCGACGCTATGGAAGTGCCTTC 837
Db 778 TATGAGATTAATAGGAATTAACAAATGGAGACTTCTATCATGCAATTTGATCGTCACT 837
QY 838 TTGATTTCTAATGCTACTCAAAATCGTACTGTTAAGCTTGAGATTGACTCAAAATCATATT 897
Db 838 TTATTAGCAGTGAAGTGGAATTAATGTTTAAATTAAGTACAGGTAATGAATTAGTT 897
QY 898 TCAGTCAAGTTAATCACTGAGGTTGTTAAGGTAACGAGGATTTAGATATTGTTAGT 957
Db 898 GAACTTTTCATCTACTCTCTGAAATTTGTTACTGTTTAAAGAAAGTTTAAACGCTAATGAT 957
QY 958 CAGTCTGCTAGTATTTAACTATCAGCTTCAATCCAACTTACCTTATTTAGTCTTTAAA 1017
Db 958 GTAGAAGCGGAAACTTGAATAATTTCTTTCAACTCAAAATCATGATGATGCTTTAAA 1017
QY 1018 GCTATTTAAAGTGAACAGTAAATAATTCATTTCTTATCACCAGTTTCGACCACTTACCCCTA 1077
Db 1018 GCCATTGATATGATGAAGTAGAGTAGAGTAGAATCTTTTGGTACAAATGAACCACTTATCTTA 1077
QY 1078 ACACAGCGGATGAGA 1094
Db 1078 AAACCAAAAGATGATGA 1094
```

RESULT 12

US-10-282-122A-35724
; Sequence 35724, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282.122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 35724

; LENGTH: 1131

; TYPE: DNA

; ORGANISM: Staphylococcus haemolyticus

US-10-282-122A-35724

Query Match 21.5%; Score 243.4; DB 16; Length 1131;

Best Local Similarity 52.8%; Pred. No. 9.9e-41;

Matches 598; Conservative 0; Mismatches 526; Indels 9; Gaps 3;

```
QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 1 ATGATGGAATTCACAAATTTAGAGAGATTTATTTTAAATCAATTAATTAATGACACATTA 60
QY 61 GCTATTAGCACTAAATGCGCATTCCTTATCTTTCATCAATAAAATGAAAGTCACTTCT 120
Db 61 GCCATCTCACAAGAACACACATTACCAATTTTAAGGGGTATCAAAATCGATGCTTAAGAT 120
QY 121 ACAGAGTAACTTTTAAACAGGGTCTAAGCGGTCAAAATATCAATTTGAAACACATTTCTCTGTA 180
Db 121 AACGAAGTCATTTCTTACTGTTTCAAGATCTCTGAGATATCTATTGAAATTTACAATCCCTAAA 180
QY 181 AGTAATGAAATGCTGGTTTCTGCTAATTAACCTCT---CCAGGAGCTATTTTATTAGAACGT 237
Db 181 CAAGTAGATGGTGAGGATATTTGTCATCTATTTCTGAAAACAGGTTTCAGTTGTACTTCTCGT 240
QY 238 AGTTTTTTTAAATATTTTCAAGTTTGGCCAGATATTAGTATAAATGTTTAAAGAAAT 297
Db 241 CGTTTCTCGTAGATATTTTAAATAAACTACACAGTAAAGATGTTAAATTTATCACAAT 300
QY 298 GAACAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGGAAAGAT 357
Db 301 GAACAATTTCAACACATGATTACTTTCAGGACATTTCTGAATTTAACTTAAGTGGTTTAGAT 360
QY 358 GTTGACCAAGTATCTCTGCTTACAAGAAGTATCAACAGAAATCTCTTGAATTTTAAACA 417
Db 361 CCTGATCAATATCCATTTATTTACCTCAAGATATCAGTGAATGCAATTTCAATTTATCAGTA 420
QY 418 AAATTTTGAAGTCTATTTATTTGCGAAACAGCTTTTTCAGCAGCTTTTACAAAGAAAGTGT 477
Db 421 AAAGTATTTAAATAATATCAATTTGCACAAACAATTTTCGCAAGTGTCCACCTCAGAAACACGC 480
```

```
QY 478 CCTATTTTAAACAGGAGTTTATTTGTTATTAAGTAATCATTAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTTACTGGTGTAACTGGCTTATA--CAAGATAATGAATTAATATATCACTGCG 537
QY 538 ACTGACTCTCATCGTATGAGCCACGTTTAAATCACTTTGGACATACTTTCAGCAGATTG 597
Db 538 ACCGATTCAACCGCTTGGCTGTGAAGAAAGTTAAATTAGAAGATGACTCTGAAACAAA 597
QY 598 ATGGTAGTCTTCCAGTAATCTTTGAGAGATTTTTCAGCAGATTTTACAGATGATATT 657
Db 598 AATGTCTATCTCCAGTAAGCTTTTATCAGAAATTAATAAATTAATGAGTGATGCGAT 657
QY 658 GAGACCGCTTGAGGTATTTTCTCACCAGCCAACTTGTTCAGAGTGAACACATTTCT 717
Db 658 GAAGAGATTGATTTTCTTCTTCAATCAAGTTTATTTAAAGTTGGAATGAAAC 717
QY 718 TTTTATACAGCTCTTAGAAGAAATTTCCCGATACAGACGCTTTTATTAATGACAGAA 777
Db 718 TTTTATTTCTCGCTTATTAAGAGACATTTCTCTGATACACACGTTTATTCCTGAGAT 777
QY 778 TTTGAGACGAGGTTTCTTCAATACCCAACTCCCTTCGCCACGCTATGGAACGTGCTTC 837
Db 778 TAGCAATTAATAGCTTTAGCAACGGTGAATTTCTATCAGCGCAATCGATCGTCTCT 837
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTGTTAAGCTTTGAGATTACTCAAAATCATATT 897
Db 838 TTATTTGGCAGGAGGTTGGCAACAACTGTTTAAAGTTAAGTACGGTAAATGATGTTGA 897
QY 898 TCAGCTCATGTTAACTCAGCTAGGTTGGTAAGTAAGTAACGAGATTTAGATATTGTTAGT 957
Db 898 GAATTAATCACTACGCTCAGCTGAGATTTGTTACTGTAAAGAAAGAAAGTTACAGCAACGAT 957
QY 958 CAGTCTGGTAGTGAATTAATCACTATCAGCTTCAATCAACTTACCTTTATGAGCTTTTAAA 1017
Db 958 GTAGAGGTTGGCACTTGAAGATTCTTTCATCTTCTTCACTCTGTTATGATGATGCGTTAAA 1017
QY 1018 GCTATTAAGTAAGTAACAGTAATAAATTCATTTCTTATCACCAGTTTCGACCAATTCACCTTA 1077
Db 1018 GCGATTGTAATGATGAGTGAAGTTGAATTTCTTTGGTACGATGAAACCGTTTCACTCT 1077
QY 1078 ACACAGGCGGATGAGGAGAAAGTTTATCCCAATTAATTAACACGATGAGAAC 1130
Db 1078 AARCCA--AAGAAGATGATTCTGTCTCAATTAATTTTACCCAATTAGAAC 1127
```

RESULT 13

```
US-09-815-242-8628
; Sequence 8628, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8628
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
; US-09-815-242-8628

Query Match      20.8%; Score 236.2; DB 9; Length 1134;
Best Local Similarity 52.2%; Pred. No. 3.1e-39;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTTCAATTTTCAATTAATCGCACATTAATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 1 ATGATGGAATTCACCTATTAAAGAGATTAATTTTATTAACAATTAATGACACATTAAAA 60
QY 61 GCTATTAGCACTAAATGCCATTCCTATTCTTTTCATCAATAAATAATGAAGTCACATTCT 120
Db 61 GCTATTTCCACCAAGAACAACTTACCTATATTAACCTGTTATCAAAATCGATGCGAAAGAA 120
QY 121 ACAGAGTAACCTTTAAACAGGGTCTAACGGTCAATATCAATTTGAACACACTATTCCTGTA 180
Db 121 CATGAAGTTATTAATTAACCTGGTTCAGACTCTGAAATTTCAATAGAAATCACTATTCCTAAA 180
QY 181 AGT---AATGAATGCTGGTTGCTTAATTAACCTCTCCAGGAGCTATTTTATAGAGCT 237
Db 181 ACTGTAGTACGGAAGATATTTGCAATATTTTCAGAAACAGGCTCAGTAGTACTTCTCTGGA 240
QY 238 AGTTTTTTTAAATATTTTTCAGAGTTTCGCCAGATATTAAGTATAATGTTAAAGAAATT 297
Db 241 CGATTTCTTTGATATTAATAAATAATTAACCTGGTAAAGATGTTAAATTAATCTACAAAT 300
QY 298 GAACAACACCAAGTTGTTTAAACAGGTGTAATCAAGATTAACCTTTTAAAGGAAAGAT 357
Db 301 GAACAATTCAGACATTAATTAACATCAGGTCAATCTGAAATTTAATTTAGTGGCTTAGAT 360
QY 358 GTTGACCAAGTATCCTCGCTACAGAGATTAACAGAAATCCTTTGATTTTAAACAA 417
Db 361 CCAGATCAATATCCTTTTATTAACCTCAAGTTTTCAGAGATGACCAATTTCAATTTGCGGTA 420
QY 418 AAATTTATGAAGTCTATTATTGCTGAAAACAGCTTTTTCAGCCAGTTTTCAGAGAAAGTCGT 477
Db 421 AAAGTCTTAAACACGTTGTTGCAACAACAAATTTTCAGTGTCCACCTCAGAAACACGC 480
QY 478 CCTATTTTAAACAGGATTCATATGTTATTAAGTAATCATTAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTAACTGGTGTGAACCTGGCTTATA---CAAGAAATGAATTAATATGACACAGCG 537
QY 538 ACTGACTCTCATCGTATGAGCCAACTTTAAATCACTTTTCGACATACTTTACAGCAGATTG 597
Db 538 ACTGACTCACCGCTTGGCTGTGAAGAAAGTTGAGTAGAGATGTTTCTGAAAAACAA 597
QY 598 ATGGTAGTCTTCCAAAGTAAATCTTTTGAAGAAATTTTCAGCAGATTTTACAGATGATATT 657
Db 598 AATGTCTATCTCCAGTAAAGCTTTTAGCTGAATTAATAAATTAATGCTGCAATGAA 657
QY 658 GAGACCGTTGAGGTATTTTCTCACCAGCCAACTTCTCAGCAAGTATTTGAGAGTGAACACATTCT 717
Db 658 GAAGACATTGATTTCTTCTGCTTCAACCAAGTTTATTTTAAAGTTGGAATGTAAC 717
QY 718 TTTTATACACGCTCTTTAGAAGAAATTTATCCCGATACAGACGCTTTTATTAATGACAGAA 777
Db 718 TTTTATTTCTCGATTATTAAGAGGACATTTCTGATACACACGTTTATTCCTGAAAC 777
QY 778 TTTGAGACGAGGTTGTTTCAATACCCNAATCCCTTCGCCACGCTATGGAACGTGCTTC 837
```


Db 778 TATGAAATTAATAAGTATAGCAATGGGAGTTTATCATGCGATTGATCGTCTCT 837
Qy 838 TTGATTTCTAATGCTACCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAAATCATATT 897
Db 838 TTATAGCGGTGAAGTGGTAAATACCTTATTAATTAAGTACAGGTGATCGTGT 897
Qy 898 TCAGCTCATGTTAACTACCTGAGGTGGTAAAGTAAACGAGGATTAAGATATTGTTAGT 957
Db 898 GAAATGCTCTTACATACCAAGAAATTTGTTACTGTAAAGAAAGAGTTGATGCAACGAT 957
Qy 958 CAGTCTGCTAGTGTATTAATCACTATCAGCTTCAATCAACTTACCTTATTTAGTCTTTAAA 1017
Db 958 GTTGAAGTGGTAGCCTGAAATTTCAATCACTCTAAATATATGATGATGCTTTAAA 1017
Qy 1018 GCTATTAAGTGAACACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCACTTACCCCTA 1077
Db 1018 GCAATCGAATATGATGAGTTGAAAGTTTCAATTTCTTCGGTACAAATGAAACCAATTTTCTA 1077
Qy 1078 ACACGAGCGGATGAGGA 1094
Db 1078 AAACCAAAAGGTGACGA 1094

RESULT 14

US-10-449-830A-1
; Sequence 1, Application US/10449830A
; Publication No. US20040137516A1
; GENERAL INFORMATION:
; APPLICANT: Phagotech Inc.
; APPLICANT: Dubow, Michael
; APPLICANT: Gros, Philippe
; APPLICANT: Bergeron, Dominique
; TITLE OF INVENTION: COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS AU
; TITLE OF INVENTION: AND ITS ENCODED PROTEIN STAAU_R2
; FILE REFERENCE: Q79017
; CURRENT APPLICATION NUMBER: US/10/449,830A
; CURRENT FILING DATE: 2003-05-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01754
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/727,892
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/885,561
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1134)
; OTHER INFORMATION:

US-10-449-830A-1

Query Match 20.8%; Score 236.2; DB 17; Length 1134;
Best Local Similarity 52.2%; Pred. No. 3.1e-39;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;
Qy 1 ATGATTCATTTTCAATTAATGCGACATATTATTTATTCATGCTTTTAAATCAACATAACGT 60
Db 1 ATGATGGAATTCACATATTAAAAGAGATTATTTATTAACAATTAATGACATTAATA 60
Qy 61 GCTATTAGCACTAAAATGCCATTCCTATTCTTTCATCAATAAAAATGAAGTCACTTCT 120
Db 61 GCTATTTCACCAAGAACACATTCCTATTATTAATGCTGATCAAAATCGATGCGAAGAA 120
Qy 121 ACAGAGTAACCTTTAAACAGGGTCAACGCTCAATATCAATTTGAAAACACTATTCTCTGA 180
Db 121 CATGAAGTATATTAACCTGTTTCACTCTGAAATTTCAATAGAAATCACTATTCTTAA 180

RESULT 15

US-08-781-986A-153
; Sequence 153, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255

Qy 181 AGT---AATGAAATGCTGGTTTCTTAATTAACCTCTCCAGGAGCTATTTTATTAGAGCT 237
Db 181 ACTGTAGATGCGAAGATATTGTCAATATTTTCAAGAACAGGCTCAGTAGTACTTCTTGA 240
Qy 238 AGTTTTTTTATTAATATTTATTTCAAGTTTGGCAGATATTAGTATAAATGTTTAAAGAAAT 297
Db 241 CGATTTCTTTGATATTTATAAAAAAATACCTGGTAAAGATGTTAAATTTATCTACAAT 300
Qy 298 GAACAACACCAAGTTGTTTTAACCCAGTGGTAAATCAGAGATTACCTTTAAAGGAAAGAT 357
Db 301 GAACAAATCCAGACATTAATTAACATCAGTCAATCTGAATTTAAATTTAAGTGGCTTAGAT 360
Qy 358 GTTGACCAAGTATCTCTGTCTACAAGAGATATACACAGAAATCTTTTGAATTTTAAAAACA 417
Db 361 CCAGATCAATATCTCTTTATTAACCTCAAGTTTCTAGAGATGACGCAATTTCAATTTGTCGTA 420
Qy 418 AAAATTTTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAAGAAAGTCT 477
Db 421 AAAGTCTTAAACAGGTGATGCAACAACAATTTTGCAGTGTCCACCTCAGAACACGC 480
Qy 478 CCTATTTTAAACAGAGTTTCATTTTGTATTAAGTAAATCATAAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTAATCTGTTGNACTGGCTTATA--CAAGAAATGAATTAATATGACACGC 537
Qy 538 ACTGACTCTCATGTATGAGCCCAAGTTTAACTACTTTGGACAATACTTTCAGCAGATTG 597
Db 538 ACTGACTCACACCCGCTTGGCTGTAAAGAAAGTTGCAAGTTAGAGATGTTTCTTGAACAAA 597
Qy 598 ATGGTAGTTCTTCAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATT 657
Db 598 AATGTCAATCATCCAGTTAAGGCTTTAGTGAATTAATAAATTTATGCTCGACANTGAA 657
Qy 658 GAGACCGTTGAGGTATTTTCTCAACCAAGCCAAATCTTTGTTTCAAGAGTGAACACATTTCT 717
Db 658 GAAGACATTTGATATCTTCTTCTTCAAAACCAAGTTTATTTTAAAGTTGGAATGTGAAC 717
Qy 718 TTTTATACACGCTCTTAGAAGGAATATCCGATACAGACCGTTTATTTATTAATGACAGAA 777
Db 718 TTTATTTCTCGATTATTAGAGAGCAATTTATCTGATACACACGTTTATTTCCCTGANAAC 777
Qy 778 TTTGAGAGCGGAGGTGTTTTCATACCAATCCCTTCCGACGCTTATGGAAGTGCCTTC 837
Db 778 TATGAATTAATTAAGTATAGACANTGGGAGTTTATCATGCAATGATGCTGCTCT 837
Qy 838 TTGATTTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATT 897
Db 838 TTATTAGCGGTGAAGTGGTAAATGAAATTTAAATTTAAATTAAGTACAGGTGATGACGCTGT 897
Qy 898 TCAGCTCATGTTAACTCACCTGAGGTGGTAAAGTGAAGGATTAAGATATTTAGTATT 957
Db 898 GAATTTGCTTCTACATCACCAAGAAATTTGGTACTGTAAAGAAAGAAAGTTGATGCAACGAT 957
Qy 958 CAGTCTGCTAGTGAATTTAACTATCATCAGCTTCAATCAACTTACCTTATTTGAGTCTTTAAA 1017
Db 958 GTTGAAGTGGTAGCCTGAAATTTTCAATCACTCTAAATATATGATGATGCTTTAAA 1017
Qy 1018 GCTATTAAGTGAACACAGTAAATTTCAATTTCTTATCACCAGTTTCGACCAATTCACCTA 1077
Db 1018 GCAATCGAATATGATGAGTTGAAATTTGAAATTTCTTCGGTACAAATGAAACCAATTTTCTA 1077
Qy 1078 ACACGAGCGGATGAGGA 1094
Db 1078 AAACCAAAAGGTGACGA 1094

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 153:
SEQUENCE CHARACTERISTICS:
LENGTH: 2347 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-153

```

```

Query Match          20.8%; Score 236.2; DB 8; Length 2347;
Best Local Similarity 52.2%; Pred. No. 4.3e-39;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTCATATTTTCAATTAATCGCACATTAATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 393 ATGATGGAATTCACATTAATAAGAGAGATTAATTTTATACACAATTAATGACACATTAATAA 452
QY 61 GCTATTAGCACATAAAATGCCATTCCTATTCCTTCATCAATAAATAATGAAGTCACTTCT 120
DB 453 GCTATTTCACCAAGAACAACTATTCCTATTAATTAATGATGATCAAAATCGATCGCAAGAA 512
QY 121 ACAGAGTAACCTTTAACAGGGTCTAACGGTCAAAATATCAATTTGAAACACATTAATTCCTGTA 180
DB 513 CATGAAGTTATATTAACCTGGTTCAGACTCTGAATTTCAATAGAAATCACTATTCCTTAA 572
QY 181 AGT---AATGAAATGCTGGTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 237
DB 573 ACTGATAGATGCGAAGATTAATGATCAATTAATTCAGAAACAGGCTCAGTAGTACTTCTTGA 632
QY 238 AGTTTTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 297
DB 633 CGATCTTTGTTGATTAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 692
QY 298 GAACAAACCAAGTTGTTTTAACAGTGGTAAATCAGAGATTAATTAATTAATTAATTAATTAATTA 357
DB 693 GAACAAATTCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 752
QY 358 GTTGACCAAGTATCTCGCTACAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 417
DB 753 CCAGATCAATATCCCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 812
QY 418 AAATTTTGAAGTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 477
DB 813 AAAGTGTCTTAAACAGCTGATTTGCAACAAATTTTGCAGTGTCCACTCAGAAACAGCG 872
QY 478 CCTATTTTAAACAGAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 537
DB 873 CCAGTACTACTGGTGAACCTGGCTTATA---CAAGAAATTAATTAATTAATTAATTAATTAATTA 929

```

```

RESULT 16
US-10-329-624-153
; Sequence 153, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997

```

```

QY 538 ACTGACTCTCATCGTATGAGCAACGTTTAACTTTGGCAATATCTTGGCAATATCTTACGACGATTG 597
DB 930 ACTGACTCACACCGCTTGGCTGTGAAGAAAGTTGCAGTTAGAGATGTTCTTGAAAAACAA 989
QY 598 ATGGTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGAGTATTTTACAGATGATATT 657
DB 990 AATGTCATCATTCAGGTAAAGCTTTAGCTGAATTAATAAATTAATCTCTGCAATGAA 1049
QY 658 GAGACCGTTGAGGTATTTTCTCACCAGCCCAATCTTTGTTCAAGTGAACACATTTCT 717
DB 1050 GAAGACATTTGATATCTTTCTTTGCTTCAACCAAGTTTATTTAAAGTTGGAATGTGAAC 1109
QY 718 TTTTATACACGCTCTTTAGAGAGAAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
DB 1110 TTTATTTCTCGATTTATAGAGGACATTTATCTGATACACACGTTTATTTCCCTGAAAC 1169
QY 778 TTTGAGACGAGGTGTTTTCATTAATACCAATCCCTTCGCCACGCTATGGAACGTCCTTC 837
DB 1170 TATGAATTAATTAAGTATATAGACAATGGGAGTTTATCATGCGATTTGATCGTCTCT 1229
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATT 897
DB 1230 TTATTAGCGGTGAAGGTGGTAAACGTTTATTAATTAATTAAGTACAGGTGATGCGTTGTT 1289
QY 898 TCAGCTCATGTTAACTCACCTGAGGTGGTAAAGTAAACGAGGATTTAGATATTGTTAGT 957
DB 1290 GAATTTGCTTCTACATCACCAGAAATTTGGTACTGTTAAAGAAAGAAAGTTGATGCAACGAT 1349
QY 958 CAGTCTGCTAGTATTTAACTTATCAGCTTCAATCCAACTTACCTTTACCTTTAGTCTTTTAAA 1017
DB 1350 GTTGAAGGTGAGCTGAAATTTCAATCACTCTAAATTAATTAATTAATTAATTAATTAATTA 1409
QY 1018 GCTATTAAGTGAACAGTAAATAATTTCAATTTATCACCAGTTCGACCAATTCACCCCTA 1077
DB 1410 GCAATCGATTAATGATGAGGTGGAATTTGAAATTTCTCGGTACAAATGAACCAATTTATTTCTA 1469
QY 1078 ACACGAGCGGATGAGGA 1094
DB 1470 AAACCAAAAGGTGACGA 1486

```

```

APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 153:
SEQUENCE CHARACTERISTICS:
LENGTH: 2347 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 153:
US-10-329-624-153

Query Match      20.8%; Score 236.2; DB 16; Length 2347;
Best Local Similarity 52.2%; Pred. No. 4.3e-39;
Matches 573; Conservative 0; Mismatches 518; Indels 6; Gaps 2;

QY 1 ATGATTCATTTTCAATTAATCGCACATTATTTATTCATGCTTTTAAATACAACTAAACGT 60
DB 393 ATGATGGAATTCACATATTAAGAGAGATTATTTTATACACATTAATTAATGACACATTAATA 452
QY 61 GCTATTAGCATAAAATGCCATTCCTATTCCTCATCAATAAAATTAAGTCACTTCT 120
DB 453 GCTATTGACCAAGAACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 512
QY 121 ACAGAGTAACCTTTAACAGGGTCTAACGGTCAATATCAATTAATTAATTAATTAATTAATTAAT 180
DB 513 CATGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 572
QY 181 AGT---ANTGAAATGCTGGTTGGTAACTACCTCTCAGGAGCTATTTTATTAGAACGT 237
DB 573 ACTGTAGATGCGGAAGATATTGTCATATTTTCAGAAACAGGCTCAGTAGTACTTCTCGGA 632
QY 238 AGTTTTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 297
DB 633 CATTTCTTTGTTGATTAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 692
QY 298 GAACAAACACCAAGTTGTTTAAACAGTGTGTAATCAGAGATTACCTTAAAGAGAAAGAT 357
DB 693 GAACAAATCCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 752
QY 358 GTTGACAGATATCTCGCTACAGAGATATCAACAGAAATCCCTTTGATTTTAAACA 417
DB 753 CCAGATCAATATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 812
QY 418 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 477
DB 813 AAATGCTTTAAACAGTGAATTCACAAATAATTTGAGTGCACCTCCAGAAACAGC 872
QY 478 CCTATTTTAAACAGGAGTTCATATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 537
DB 873 CCAGTACTACTGTTGGTGAAGCTTATA---CAAGAAATGAATTAATTAATTAATTAATTAATTAAT 929
QY 538 ACTGACTCTCATGTAAGGCAAGCTTTAATCACTTTGGCAATTAATTAATTAATTAATTAATTAAT 597
DB 930 ACTGACTCACACCGCTGCTGTAAGAAAGTTGACGTTAGAGATGTTTCTGAAACAAA 989
QY 598 ATGTGAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTACAGAGTATTACAGATGATATT 657
DB 990 AATGATCATCTCCAGGTAAGGCTTTAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1049
QY 658 GAGACCGTTGAGGATTTTCTCACCAGCCAAATCTTGTTCAGAGAGTGAACACATTTCT 717
DB 1050 GAAGACATTTGATATCTTCTTGTCTCAACCAAGTTTATTTAAAGTTGGAATGTGAAC 1109
QY 718 TTTTATACAGCCTCTTAGAAGAAATTAATCCCGATACAGACCGTTTATTAATGACAGAA 777

```

```

DB 1110 TTTATTTCTCGATTATTAGAAAGGACATTATCTCTGATACACACGTTTATTCCTCGAAAC 1169
QY 778 TTTGAGACGGAGGTTGTTTCAATACCCCAATCCCTTCCGCCACGCTATCGAAACGTCCTTC 837
DB 1170 TATGAATTAATTAAGTATAGACAAATGGGAGTTTATCATGCGATTGATCGCTCT 1229
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTAAAGCTTGAGATTACTCAAAATCATATT 897
DB 1230 TTTATAGCGCTGAAGGTGTAATAACCGTTATTAATTAATTAAGTACAGGTGATGACGTGTT 1289
QY 898 TCAGCTCATGTAACCTGAGGTTGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 957
DB 1290 GAATTTGCTTCTACATCACCAAAATTTGTTACTGTAAAAGAAAGAGTTGATGCAACCGAT 1349
QY 958 CAGTCTGCTAGTGAATTTAACTATCAGCTTTCATCACTCACTTATTCCTTATTCAGTCTTTAA 1017
DB 1350 GTTGAAGGTGCTAGCTGAAATTTTCATTCATCTAATATATATGATGATGATGATGATGATGAT 1409
QY 1018 GCTATTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 1077
DB 1410 GCAATCGAATGATGAGGTTGAAGTTGAAATTCCTTCGCTAGCAATGAACCAATTTATCTTA 1469
QY 1078 ACACGAGCGATGAGGA 1094
DB 1470 AAACCAAAAGGTGACGA 1486

```

RESULT 17

```

US-09-815-242-8602
; Sequence 8602, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8602
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
US-09-815-242-8602

```

Query Match 20.5%; Score 233; DB 9; Length 1134;
Best Local Similarity 52.1%; Pred. No. 1.5e-38;

	Matches	571;	Conservative	0;	Mismatches	520;	Indels	6;	Gaps	2;
QY	1	ATGATTCAATTTT	CAATTAATCGC	ATTATTTATTTATTCATGCTTTTAAATACAAC	TAAAGCT	60				
DB	1	ATGATGGAATTC	TATTAATAAAGAGAT	TATTTATTAACAATTAATTCACACATTTAAA	60					
QY	61	GCTATTAGACCT	TAAAAATGCG	ATTCTTATCTTTTCATCAATAAAATTTGAAGTCACTTCT	120					
DB	61	GCTATTTCAC	CAAGAACACAT	TACTCTATTTAACTGGTATTTAAATCGATGCAAAAGAA	120					
QY	121	ACAGGAGTAAC	TTTAAACAGGGT	CTAACCGTCAAAATATCAATTTGAAACACACTATTTCCCTGTA	180					
DB	121	CATGAAGTTAT	TATTAAC	TGGTTCAGACTCTGAAATTTCAATAGAAATCATCTATTTCCCTAAA	180					
QY	181	AGT---	AATGAAATTC	GTGGTTCGTAAATACCTCTCCAGGAGCTATTTTATTAGAAGCT	237					
DB	181	ACTGTAGATGG	CGGAAGATAT	TGTCAATATTTTCAGAAACAGGCTCAGTAGTACTTTCCTGGA	240					
QY	238	AGTTTTTTTT	TATTAATTTT	CAAGTTTGGCAGATATTAGTATAAATGTTAAAGAAATTT	297					
DB	241	CGATTCCTT	TGTTGATTTAT	TAAAAAATTTACTGTTAAGATGTTAAATATATCTACAAAT	300					
QY	298	GAAACAACA	CCAGTTGTTTTT	AACAGTGGTAAATCAGAGATTAACCTTTAAAGGAAAGAT	357					
DB	301	GAAACAATTC	CGACACATTAAT	TACATCAGETCATTTCTGAAATTTAATTTAAGTGGCTTAGAT	360					
QY	358	GTTGACCAGTAT	CTCTCGTTC	ACAAGAGTATCAACAGAAAATCCTTTGATTTTAAAAACA	417					
DB	361	CCAGATCAAT	TATCTTTT	TATACCTTCAAGTTTCTAGAGATGCGCAATTCAAATTTGTCGGTA	420					
QY	418	AAATTTTGA	AGTCTATPAT	TGCTGAAACAGCTTTTTCAGCCAGTTTTCACAAGAAAGTCGT	477					
DB	421	AAAGTGCTT	AAAAACGTAAT	TGCAACAAAGAAATTTTCAGTGTCCACCTCAGAAAACAGC	480					
QY	478	CCTATTTTAA	CAGGAGTTCAT	TGTATTAAGTAAATCATAAAGATTTTAAAGCAGTAGCG	537					
DB	481	COAGTACT	TA	CTGGTGGAACTGGCTTTATA--CAAGAAAATGAAATTAATATGCACAGCG	537					
QY	538	ACTGACTCT	CTCATCTAT	GAGCCAAAGCTTTTAACTACTTTTGACAAATCTTCAGCAGAGATTTG	597					
DB	538	ACTGACTCAC	ACCGCTTGG	CTGTAAAGAAAGTTGCAGTTAGAAGATGTTCTGAAAACAAA	597					
QY	598	ATGGTAGTCT	TCCAAATTAAT	CTTTGAGAGAAATTTTCAGCAGTATTTACAGATGATAT	657					
DB	598	AATGTCTAT	CAATCCAGGTAA	AGGCTTTAGCTGAAATTAATAAAATATATGTCGACAATGAA	657					
QY	658	GAGACCGT	TCAGGTATTTTT	CTCACCAAGCCAAATCTTTGTTTCAGAAGTGAACACATTTCT	717					
DB	658	GAAGACAT	TGATATCTTCT	TGCTTCAACCAAGTTTATTTAAAGTTGGAATGTGAAC	717					
QY	718	TTTTATAC	CGCTCTTAGA	AGGAATATATCCCGATACAGACCGTTTATTAATGACAGAA	777					
DB	718	TTTTATTTCT	CGATTTATTA	GAGACATATCTCTGATACAACACAGCTTTATTCCTCGTAAAC	777					
QY	778	TTTGAGAC	GGAGGTTGTTTT	CAATACCCAATCCCTTCGCCACGCTATGNAAGTGCCCTTC	837					
DB	778	TATGAATTA	ATAATTAAGTAT	TAGCAATAGGGAGTTTATCATCGCAATGATCGTGCATCT	837					
QY	838	TTGATTTCT	TAATGCTACT	CAAAATGGTACTGTTTAAAGCTTTGAGATTACTCAAAATCATAT	897					
DB	838	TTATTAGC	ACGTGAAGGTGGT	ATAATACGTTATTTAAATTAAGTACAGGTGATGACGTTGTT	897					
QY	898	TCAGCTCAT	GTTAACTCAC	CTGAGGTTGGTAAACGAGGATTTAGATATTTGTTAGT	957					
DB	898	GAATATCT	TCTTACATC	ACCAGAAATTTGGTACTGTATAAAGAGAAAGTTGATGCAAAACGAT	957					
QY	958	CAGTCTGGT	TAGTGAITTA	ATCATCTCAGCTTCAATPCCAACTTACCTTATTTGAGTCTTTAAAA	1017					
DB	958	GTTGAAAG	GGTGTAG	CTCGTAAAAATTTCAATCAACTCTAAATATATGATGATGCTTTAAA	1017					
QY	1018	GCTATTAA	AGTGAACACAGT	TAATAATTCATTTCTTATCACCGAGTTCGACCAATTCACCTA	1077					
DB	1018	GCAATCGA	TAAATGAGGGT	TGAAGTTGAATTTCTTCGGTACAAATGAACCAATTTATCTGA	1077					

```

QY      1078 ACACGAGCGATGAGGA 1094
      | | | | |
Db      1078 AAACCAAAAGGTGACGA 1094

RESULT 18
US-10-282-287-9
; Sequence 9, Application US/10282287
; Publication No. US20030129633A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Zhang, Dan
; APPLICANT: Whipple, Richard
; TITLE OF INVENTION: DNA REPLICATION PROTEINS OF GRAM POSITIVE BACTERIA AND
; TITLE OF INVENTION: THEIR USE TO SCREEN FOR CHEMICAL INHIBITORS
; FILE REFERENCE: 22221/1002
; CURRENT APPLICATION NUMBER: US/10/282,287
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/235,245
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/074,522
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/093,727
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-282-287-9

```

Query Match	20.5%;	Score 233;	DB 15;	Length 1134;
Best Local Similarity	52.1%;	Pred. No. 1.5e-38;		
Matches 571;	Conservative 0;	Mismatches 520;	Indels 6;	Gaps 2
QY	1	ATGATTCAATTTTCAATTAATTCGCACATATTTATTCATGCTTTAAATACAACTAAACGT	60	
DB	1	ATGATGGAAATTCACATTTAAAGAGATTTATTTATACAAATTAATGACACATTAATAA	60	
QY	61	GCTATTAGCACATAAAATGCCATTTCCTTATCTCTTCATCAATAAAAAATTTGAAGTCACITCT	120	
DB	61	GCTATTTCCACCAAGACAACATTACCTATATTAAGTATCAAAATCGATGCGAAGAA	120	
QY	121	ACAGGAGTAACCTTTAAACAGGGCTTAACGGTCAAAATATCAAAATGAAAAACACTATTCCTGTA	180	
DB	121	CATGAAGTTATATTAACATCTGGTTCAGACTCTGAAATTTCAATAGAAATCCTATTCCCTAAA	180	
QY	181	AGT---AATGAAAAATGCTGGTTGCTAATACCTCTCCAGGAGCTATTTATTATAGAAGCT	237	
DB	181	ACTGTAGATGCGAAGATTTGTCAATATTTTCAGAAACAGGCTCAGTAGTACTTCCTCGGA	240	
QY	238	AGTTTTTTTATTAATATTTATTCAGGTTTGCAGATATTAGTATAAAATGTTTAAAGAAATT	297	
DB	241	CGATTCCTTTGTGATATATAAAAAAATTAACCTGGTAAAGATGTTAAATATCTACAAAT	300	
QY	298	GAACAAACCAAGTTGTTTTAAACAGTGGTAAATACAGAGATTACCTTTAAAGGAAAAAGAT	357	
DB	301	GAACAATTCAGACATTAATATACATCAGGTCAATTCGAAATTTAATTTGAGTGGCTTAGAT	360	
QY	358	GTTGACCAGGTATCCTCGCTACAGAAGTATCAACAGAAAATCCTTTGATTTTAAAAACA	417	
DB	361	CCAGATCAATATCCTTTATTAACCTCAAGTTCTAGAGATGCGCAATTCCAATTTGTCGGTA	420	
QY	418	AAATTAATGAGTCTATTATTTGCTGAAACAGCTTTTTGCAGCCAGTTTTCACAGAAAGTCGT	477	
DB	421	AAAGTACTTAAAAACGATTTGCACAAACGAAATTTTGCAGTGTCCACCTCAGAAACACGC	480	
QY	478	CCTATTTTACAGAGTTTCATATTTGATTTAAGTTAATCATAAAGATTTTAAAGCAGTACGC	537	
DB	481	CCAGTACTACTCGTGTGAATCGGCTTATA--CAAGAAATGAATTAATATGCACAGCG	537	

```

QY 538 ACTGACTCTCATGATGAGCAACGTTTAAATCACTTTGGACAATACTTTCAGCAGATTG 597
Db 538 ACTGATTCACACCGCTGCTGTAAGAAAGTTGCGAGTTAGAAAGATGTTTCTGAAACAAA 597
QY 598 ATGCTAGTCTTCCAAAGTAAATCTTTTGAGAGAAATTTTCAGCAGATTTTACAGATCATATT 657
Db 598 ATGCTATCATCTCCAGGTAGGCTTTAGCTGAATTAATAAATTAATGTTCTGCAATGAA 657
QY 658 GAGACCGTTGAGGTATTTTCTCACCAAGCCAAATCTTTGTCAGAAAGTGAACACATTTCT 717
Db 658 GAAGACATGATATCTCTTCTGCTTCAAACCAAGTTTATTTAAAGTTGGAATGTGAAC 717
QY 718 TTTTATACAGCCTCTTGAAGGAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
Db 718 TTTATCTCGATTTATGAAGGACATTTATCTGATACAAACGTTTATTCCTCGAAAC 777
QY 778 TTTGAGCGGAGTGTCTTCAATACCAATCCCTTCGCCACGCTATGGAAGTCGCTTC 837
Db 778 TATGAATTAATTAAGTATAGACAAATGGGAGTTTATCATCGATGATGTCGCTCT 837
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTAAAGCTTGAGATTAATCAAAATCATATT 897
Db 838 TTTATAGCAGCGTGAAGTGGTAAATACGTTATTAATTAAGTACAGGTGATGACGTTGTT 897
QY 898 TCAGCTCATGTTAACTCAGCTGAGGTGGTAAAGTAAACGAGGATTTAGATATTGTTAGT 957
Db 898 GAATATATCTTACATACCAAGAAATTTGTTACTGTAAAGAAAGAAAGTTGATGCAAAACGAT 957
QY 958 CAGTCTGTAGTGAATTAACATACAGCTTCAATCAACCTTACCTTATTTAGTCTTTAAAA 1017
Db 958 GTTGAAGTGTGAGCTGAAATTTCAATCAACTCTAATATATGATGATGCTTTAAAA 1017
QY 1018 GCTATTAAGTGAACAGTAAATTCATTTCTTATCACCGATTCGACCAATTCACCTTA 1077
Db 1018 GCAATCGATATGATGAGTGTGAAGTGTGAATCTTTCGTTACAAATGAACCAATTTATCTA 1077
QY 1078 ACACGAGCGATGAGGA 1094
Db 1078 AAACCAAAAGGTGACGA 1094

```

RESULT 19

```

US-10-282-122A-7900
; Sequence 7900, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Walli, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7900
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-282-122A-7900

Query Match 20.5%; Score 233; DB 16; Length 1134;
Best Local Similarity 52.1%; Pred. No. 1.5e-38;
Matches 571; Conservative 0; Mismatches 520; Indels 6; Gaps 2;

QY 1 ATGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGT 60
Db 1 ATGATGGAATTCATTTTAAAGAGATTTATTTATACAAATTAATTAATGACACATTAAA 60
QY 61 GCTATTAGCAGTCAAAATGCCATTTCTTCTTTCATCAATAAAATTTGAAGTCACTTCT 120
Db 61 GCTATTTCCACCAAGAACAACTTACCTATATTAACGTGATCAAAATCGATCGAAGAA 120
QY 121 ACAGAGTAACTTTTAAAGGCTTAAACGGTCAATATCAATTTGAAAACACATTTCTGT 180
Db 121 CATGAAGTTATCTAACTGGTTTCAAGCTCTGAAATTTCAATGAAATCACTATTTCTAAA 180
QY 181 AGT---AATGAAATGCTGGTTTGTAAATTAACCTCCAGGAGCTATTTTATTAGAGCT 237
Db 181 ACTGATGATGCGAAGATTTGTCAATATTTCAAGAACAGGCTAGTAGTACTTCTTGA 240
QY 238 AGTTTTTTTAAATTAATTTTCAAGTTTGGCAGATATTAGTATTAATTTTAAAGAAAT 297
Db 241 CGATTTCTTTGATATTTATAAAAAATTAACCTGCTGAAAGATGTTTAAATTTATCTCAAA 300
QY 298 GAACAAACCAAGTTGTTTAAACCGTGTAAATCAGAGATTTACCTTTAAAGGAAAGAT 357
Db 301 GAACAAATTCAGACATTTAAATTAATCACTAGCTCATTTCTGAATTTTAAATTTAAGTGGCTTAGAT 360
QY 358 GTTGACAGATATCTCTCGTCTCAAGAGATATCAACAGAAATCTCTTGAATTTTAAACAA 417
Db 361 CCAGATCAATATCTTTTATTAACCTCAAGTTCTAGAGATGACGCAATTCATTTGCGGT 420
QY 418 AAATTTTGAAGTCTATTTATTTGCTGAAACAGCTTTTGCAGCCAGTTTACAAGAAAGTCT 477
Db 421 AAAGTGTCTTAAACAGCTGATTTGCACAAACGAAATTTTGCAGTGTCCACCTCAGAAACACGC 480
QY 478 CCTATTTTAAACAGGATTCATATTTGTTATTAAGTATCAATTAAGATTTTAAAGCAGTAGCG 537
Db 481 CCAGTACTAATCTGTTGTAAGTGTGCTTATA---CAAGAAATTAATTAATGATGACAGCG 537
QY 538 ACTGACTCTCATGATGAGCAACGTTTAAATCACTTTTGGACAATACTTTCAGCAGATTG 597
Db 538 ACTGACTCAGACCGCTTGGCTGTGAAGAGTTGCGATTTAGAGATGTTTCTGAAACAAA 597
QY 598 ATGCTAGTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTTACAGATGATATT 657
Db 598 AATGTCATCAATTCAGGTGAAGCTTTAGCTGAATTTAAATTAATTAATGCTGCAATGAA 657
QY * 658 GAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAAAGTGAACACATTTCT 717
Db 658 GAAGACATTTGATATCTCTTCTTCTTCAACCAAGTTTATTTTAAAGTTGGAATGTGAAC 717
QY 718 TTTTATACACGCTCTTGAAGGAATTTATCCGATACAGACCGTTTATTAATGACAGAA 777
Db 718 TTTATTTCTGATTTAATTAAGAGGACATTTATCTGATACAAACGTTTATTTCTCCTGAAAC 777

```

QY 778 TTTGAGACGAGGTGTTTTCATATACCAATCCCTTCGACGCTATGGAACGTGCCTTC 837
Db |||||
QY 778 TATGAATTAATTAAGTATAGACATGGGAGTTTATCATCGGATTCATCGTGCCTCT 837
Db |||||
QY 838 TTGAATTTCTAATGCTACTCAAAATGGTACTGTGTTAAGCTTGAGATTACTCAAAATCATATT 897
Db |||||
QY 838 TTATTAGCAGCTGAAGGTGGTAAATACGTTTATTAAATTAAGTACAGGTGATGACGTTGTT 897
QY 898 TCAGCTCATGTTAACTCACTGAGGTTGGTAAGGTAACGAGGATTAAGATTTGTTAGT 957
Db |||||
QY 898 GAATTAATCTTCTACATCACAGAAATGGTACTGTAAAGAGAGAGTTCATGCAACAGAT 957
QY 958 CAGTCTGGTAGTGAATTAATCACTACAGCTTCAATCCAACTTACCTTATTCAGTCTTTAAA 1017
Db |||||
QY 958 GTTGAAGGTGGTAGCTGAAATTTCAATCACTCTAAATATATGATGATGCTTTAAA 1017
QY 1018 GCTATTAAAGTGAACAGTAAATTCATTTCTTATACCAAGTTTCGACCAATTCACCCCTA 1077
Db |||||
QY 1018 GCAATCGAATGATGAGGTTGAAGTTGAATCTTCGGTACAATGAACCAATTTATTCTA 1077
QY 1078 ACACGAGCGATGAGGA 1094
Db |||||
QY 1078 AACCAAAAGGTGACGA 1094
Db |||||

RESULT 20

US-09-815-242-4614
; Sequence 4614, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4614
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4614

Query Match 20.3%; Score 230; DB 9; Length 1128;
Best Local Similarity 51.9%; Pred No. 6.1e-38;
Matches 568; Conservative 0; Mismatches 520; Indels 6; Gaps 2;
QY 4 ATTCAATTTTCAATTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTTAAACGTGCT 63
Db |||||
1 ATGGAATTCACATTTAAAGAGATTTATTTATTACACAATTAATGACACATTAAAGCT 60

QY 64 ATTAGCACATAAAATGCGCATTCCTATTCTTTATCAATAAAAAATGAAAGTCACCTTCTACA 123
Db |||||
QY 61 ATTTCCACCAAGAACCAACATTTACCTATATACTTAACTGGTATTAAATCGATGCAAAAGACAT 120
QY 124 GGAGTAACTTTAAACAGGGTCTAAACGGTCAAAATATCAATTTGAAACACATTTCTCTGTAAGT 183
Db |||||
QY 121 GAAGTTATTAATTAACCTGGTTCAGACTCTGAAAATTTCAATAGAAATCCTATTCTTAAAAACT 180
QY 184 ---AATGAAAATGCTGGTTGCTAATACCTCTCCAGGAGCTATTTTATTAGAGCTAGT 240
Db |||||
QY 181 GTAGATGCGGAAGATATTGTCAATATTTCAAAACAGGCTCAGTAGTACTTCTTGGAGCA 240
QY 241 TTTTATTATAATATTATTTCAAGTTGCGGATATTAGTATATAATGTTTAAAGAAATTTGAA 300
Db |||||
QY 241 TTTCTTTGTTGATATTATAAAAAAATTAACCTGGTAAAGATGTTAAATATTATCTACAAATGAA 300
QY 301 CAACACCAAGTTGTTTAAACAGGTTGTAATACAGATTTACCTTTAAAGGAAAGATGTT 360
Db |||||
QY 301 CAATTCAGACATTAATTACATCAGGTCATTCTGAAATTTAATTTAAGTGGCTTAGATCCA 360
QY 361 GACCAAGTATCTCGTCTACAGAAAGTATCAACAGAAATCCCTTTGATTTTAAAAACAAAA 420
Db |||||
QY 361 GATCAATATCTTTATTACCTCAAGTTTCTAGAGATGACGCAATTCATTTCTCGTAAAA 420
QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTTCAGCCAGTTTACAGAAAGTCGTCT 480
Db |||||
QY 421 GTGCTTAAAAACGTAATTTGCACAAACGAATTTTTCAGTGTCCACCTCAGAAACACGCCCA 480
QY 481 ATTTTAAACAGGATTCATATTGTTAAGTAATCATATAAGATTTTAAAGCAGTAGTACCACT 540
Db |||||
QY 481 GTACTAATGTTGTAAGTGGCTTATA--CAAGAAATGAATTAATATGACAGCCACT 537
QY 541 GACTCTCATGATGAGCCAAAGTTTAACTTTGACAAATACTTTCAGCAGATTTGATG 600
Db |||||
QY 538 GACTCACACCGCTTGGCTGTAAGAAAGTTGAGTTAGAGATGTTTCTGAAAAACAAAAAT 597
QY 601 GTAGTCTTCCAGTAATCTTTGAGAGAAATTTTCAGCAGTATTTCAGATGATATTGAG 660
Db |||||
QY 598 GTCAATCAATTCAGTAAGGCTTTAGCTGAAATTAATTAATTAATGTCGCAATGAAAGAA 657
QY 661 ACCGTTGAGGTATTTTCTCACCAGCCAAATCTGTTGAGAAAGTGAACACATTTCTTTT 720
Db |||||
QY 658 GACATTCATATCTTTTCTTCTTCAAAACCAAGTTTATTAAAGTTGAAATGTCACCTTT 717
QY 721 TATACAGCCTCTTAGAGGAAATTTATCCGATACAGCCGTTTATTATTAATGACAGAAATTT 780
Db |||||
QY 718 ATTTCTCGATTATTAGAAAGGACATTATCTCGATACAAACAGCTTTATTTCCCTGAAAACTAT 777
QY 781 GAGACGGAGGTGTTTTCATATCCCAATCCCTTCCGACGCTATGGAACGTGCTTCTTG 840
Db |||||
QY 778 GAAATTAATTAAGTATAGACATGGGAGTTTATCATGCGATTCGATTCGATCTTTA 837
QY 841 ATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTTGAGATTACTCAAAATCATATTTC 900
Db |||||
QY 838 TTAGCAGCTGAAGGTGGTAAATACGTTATTAAATTAAGTACAGGTGATGACGTTGTTGAA 897
QY 901 GCTCATGTTAACTCACCTGAGGTTGGTAAAGTAAAGAGGATTTAGATATTGTTAGTCAG 960
Db |||||
QY 898 TTATCTTCTACATCACCAAGAAATTTGATCTGTTAAAGAGAGAGTTCATGCAACGATGTT 957
QY 961 TCTGGTAGTGAATTAATCATATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTTAAAGCT 1020
Db |||||
QY 958 GAAGGTGGTAGCTGAAATTTTCATTCACCTCTAAATATATGATGATGCTTTTAAAGCA 1017
QY 1021 ATTAAAGTGAACAGTAAATTTCAATTTATCACAGGTCGACCAATTCACCTTAAACA 1080
Db |||||
QY 1018 ATCGAATATGATGAGGTTGAAGTTGAATTTCTTGGTACAATGAACCAATTTATTCTAAAA 1077
QY 1081 CCAGGCGATGAGGA 1094
Db |||||
QY 1078 CCAAAAGGTGACGA 1091
Db |||||

```
RESULT 21
US-10-398-221-3763
; Sequence 3763, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3763
; LENGTH: 4736
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; NAME/KEY: misc feature
; FEATURE:
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3763

Query Match      17.7%; Score 200.4; DB 16; Length 4736;
Best Local Similarity 51.5%; Pred. No. 1.7e-31;
Matches 565; Conservative 0; Mismatches 521; Indels 12; Gaps 4;

QY 2 TGATTCAAATTTTCAATTAATCCACATTAATTTATTCATGCTTTAAATACAACTAAACGTG 61
DB 1218 TCATGAAATTTGTTATGAGCGTGTGCTTGTCCAGCAGTCAATGAAGTTACTCGTG 1277
QY 62 CTATTAGCACTAAATGCCATTCCTATCTTTCATCATATAAAATTTGAAGTCACTTCTA 121
DB 1278 CCATCTCTGCAAGAACAGATTCCTCAATTTTAAACGGGGATAAAATAGTCGTAATGATG 1337
QY 122 CAGGAGTAATTTTAAACAGGCTCTAAGGTCCTAATTAATCAATCAAAACACTATTCCTGTAA 181
DB 1338 AAGGTGTACACTAATCTGCTAGTGTATCGGATATTTCAATCGNAGCATTTATTCATTA 1397
QY 182 GT---AATCAAAATCTGTTGCTGTAATTAATCACTCTCCAGGAGCTATTTTATTAAGCTA 238
DB 1398 TTGAAATGATCAAGTAATTTAGAAAGTGAGAGTTTGGTGGAAATTTGACTTCAATCAA 1457
QY 239 GTTTTATTAATATTTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAAATG 298
DB 1458 AATACCTTTGGCGATATTTGCTGCTTTACCAAGAAATAGTAAATTTGAAGTGACTT 1517
QY 299 AACACACCAAGTTGTTTAAACAGTGGTAATCAAGATTAACCTTTAAAGGAAAGATG 358
DB 1518 CTAACCTACCAACCAACATAGTTCTGGCCAGCATCTTTTACATTAATGCTTAGATC 1577
QY 359 TTGACCAGTATCTCGCTCAAGAGATATCAACAGAAATCTTTGATTTTAAACAA 418
DB 1578 CAATGGAATATCTTAAATTTACCTGAAGTAACAGCGGAAACAAATTAATTTCAATTA 1637
QY 419 AATTAATGAAGTCTATTATGCTGAAACAGCTTTTGCAGCAGTTTACAGAAAGTCTG 478
DB 1638 ATGTACTTTAAATAATTTGTAGACAACTGTTTGTGCTGTGCTGCGATTTGAAGTTCGTC 1697
QY 479 CTATTTTAAACAGGATTCATATTTGATTAAGTAATCAATAAGATTTTAAAGCAGTAGCGA 538
DB 1698 GAGTACTTACTGTTGTAATCTGATTAATCAAGAAATTA---CTAAGCGAGTTGCA 1754
QY 539 CTGACTCTCATGATGAGCCACGTTTAACTCACTTTGGACA---ATACTCAGCAGATT 595
DB 1755 CCGATAGTATCATGCTAGCTTTAGCGTGAATACCTCTTGAAACAGACATTTGATGAAGAT 1814
QY 596 TGATGCTAGTTCTTCCAGTAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATA 655

RESULT 22
US-10-398-221-9/c
; Sequence 9, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 684707
; TYPE: DNA
; ORGANISM: Listeria innocua
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-9

Query Match      17.5%; Score 198.8; DB 16; Length 684707;
Best Local Similarity 51.4%; Pred. No. 3e-30;
Matches 564; Conservative 0; Mismatches 522; Indels 12; Gaps 4;

QY 2 TGATTCAAATTTTCAATTAATCGCATTAATTTATTCATGCTTTAAATACAACTAAACGTG 61
DB 267740 TCATGAAATTTGTTATTTGAGCGTGTGCTTTGTCCAGCAGTCAATGAAGTTACTCGTG 267681
QY 62 CTATTAGCACTAAATAATGCCATTCCTATCTTTTCATCAATAAAATTTGAAGTCACTTCTA 121
```

Db 267680 CCATCTCTCAGACACAAACGATTCCTCAATTTTAAACGGGATAAAAATAGTCGTAATATGATG 267621
 QY 122 CAGGAGTAACCTTAAACAGGCTCTAACGGTCAATATCAATATGAAACACACTATTCCTGTAA 181
 Db 267620 AAGGTGTACTTTTAAACGGTAGGATTCGTGATATTTCCATCGAAGCAATTTATTCCTATTA 267561
 QY 182 ---GTAATGAAATGCTGGTTGCTAATACCTCTCCAGGAGCTATTTTATTAGAAGCTA 238
 Db 267560 TCGAAATGATGAAGTAATGTAGAGTTTGAAGTTTGGTGTATCGTCTTCAATCCA 267501
 QY 239 GTTTTATTATTAATATTTTCAAGTTTCCAGATATATAGTATAAATGTTAAAGAAATG 298
 Db 267500 AGTACTTCGGTGATATGTTGCTGCTTTACCTGAAGAAATGCGAAATTTGAATTTACTA 267441
 QY 299 AACCAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGCAAGAGATG 358
 Db 267440 CAAATTTACAACTACATTAAGTTCTGGCCAGCTTCTTTTACACTAACCGCTTAGATC 267381
 QY 359 TTGACAGATATCCTGCTTACAGAGATATCAACAGAAATCCTTTTGATTTTAAACAA 418
 Db 267380 CAATGGAATACCTTAAATTACCAGAGATTACTGATGGTAAATAATTAATAATTCCTATTA 267321
 QY 419 AATTATTGAAGTCTATTATTTCAAGTTTCCAGATATATAGTATAAATGTTAAAGAAATG 478
 Db 267320 ATGCTCTTAAACAACTATTAGACAAACTGTTTTCGCTGCTATCGAAGTTGCGC 267261
 QY 479 CTATTTTAAACAGGAGTTCTATTTGATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGA 538
 Db 267260 CAGTCTTACAGGTTTAACTGGATTTAT---TAAAGATAACAACTTAGCGAGTTGCA 267204
 QY 539 CTGACTCTCATGATGAGCCAACTTTTAACTCACTTTTGA---CAATACTTTCAGCAGATT 595
 Db 267203 CCGATAGTCACTGCTAGCATTTACGTGAATTTCCACTTTGAAACAAATATTGACGAAGAT 267144
 QY 596 TGATGATGTTCTTCCAGTAATCTTTGAGAGATTTTTCAGCAGTATTTACAGATGATA 655
 Db 267143 ACAACATTTGTTATTCGAGAAAGAGCTTAGCTGAACATAAATAAATTTTACAGCAGTCA 267084
 QY 656 TTGAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATTT 715
 Db 267083 GTGAATCCATTGAATGACACTCGCAATATCAATCTTATTTAAATTAAGATTTAT 267024
 QY 716 CTTTATTATACAGCCCTTTAGAGAAATTTATCCGATACAGACGTTTATTAATGACAG 775
 Db 267023 TATTTTATCTGTTTACTGGAAGGTAGTTTACCAGATACATCACGCTTAATTTCAACTG 266964
 QY 776 AATTGAGACGAGGTGTTTTCATACCCCAATCCCTTCGCCAGCTATGAAAGTGCCT 835
 Db 266963 ATACTAAATCAGAACTTGTCAATTAATCCAGAGCAATTTTAAAGCGATTGACCGTGTCT 266904
 QY 836 TCTGATTTCTAATGCTACTCAAAATGGTACTGTTTAACTTTGAGATTACTCAAAAT---C 892
 Db 266903 CCTACTTGTCTCGGAAACCGTAATACGTTTATTAATTAATGACCTTGAATGGAC 266844
 QY 893 ATATTTCAGCTCATGTTTAACTTCACTGAGTTGTTGTAAGTAAACAGAGATTTAGATTTG 952
 Db 266843 AAGTTGAAGTATCCTTCAACTCTCTGAGTTGGAATGTTTTCGGAATAATGCTTTTATGTC 266784
 QY 953 TTAGTCACTGTTGATGATTAACTATCACTCAGCTTCAATCCAACTTACCTTATTCAGTCTT 1012
 Db 266783 AAAGCTTTACTGTTGAAGAAATTAATATATCTTTTACGGTAAATATCATGATGATGAT 266724
 QY 1013 TAAAGCTATTAAAGTGAACAGTAAATAATTTCTTTATATCCAGTTTCGACCATTTCA 1072
 Db 266723 TCGCGCATTTTGAAGGCGATGATATTCAAATCTCTCTCTGGCACAATGAGACCTTTTCG 266664
 QY 1073 CCCTAACACACAGGCGATG 1090
 Db 266663 TACTTCGACCAAAAGATG 266646

RESULT 23

US-10-398-221-2058

; Sequence 2058, Application US/10398221
 ; Publication No. US20040018514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KUNST, Frederik
 ; APPLICANT: GLASER, Philippe
 ; TITLE OF INVENTION: Listeria innocua, genome and applications
 ; FILE REFERENCE: 344 702 - US
 ; CURRENT APPLICATION NUMBER: US/10/398,221
 ; PRIOR FILING DATE: 2003-03-27
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
 ; PRIOR FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: FR 00/12 697
 ; NUMBER OF SEQ ID NOS: 4025
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 2058
 ; LENGTH: 3011208
 ; TYPE: DNA
 ; ORGANISM: Listeria innocua
 ; US-10-398-221-2058

Query Match 17.5%; Score 198.8; DB 16; Length 3011208;
 Best Local Similarity 51.4%; Pred. No. 5.6e-30;
 Matches 564; Conservative 0; Mismatches 522; Indels 12; Gaps 4;

QY 2 TGATTCATTTTCAATTAATCGCACATTTATTTATTCATGCTTTAAATCAACTAAACGTG 61
 Db 1866 TCATGAATTTGTTTATTTAGCGGTGATCGTCTTGCCAAAGCAGTCAATGAAGTTACTCGTG 1925
 QY 62 CTATTAGCACTTAAATAATGCCATTCCTATTCTTTTCATCAATAAAAAATTTGAAGTCACTTCTA 121
 Db 1926 CCACTCTGCAAGAACACGATTTCCAATTTTAAACGGGATTAATAATAGTCGTAATATG 1985
 QY 122 CAGGAGTAACCTTTAAACAGGCTTAAACGGTCAAAATATCAATTTGAAACACACTATTCCTGTAA 181
 Db 1986 AAGGTGTTACTTTAAACCGTAGCGATTTCTGATATTTCCATCGAAGCATTTTATTCATTA 2045
 QY 182 ---GTAATGAAATGCTGGTTTGTCTAATTAACCTCTCCAGGAGCTATTTTATAGAACCTA 238
 Db 2046 TCGAAATGATGAAGTAATTTGTAGAGTTGAAAGTTTGGTGGTATCGTCTTCAATCCA 2105
 QY 239 GTTTTTTATTAATTAATTTTCAAGTTTGGCAGATATTTAGTATAAATGTTTAAAGAAATG 298
 Db 2106 AGTACTTCGGTGATATTTGTTGCTGTTTACCTGAAGAAATGTCGAAATTTGAAGTTACTA 2165
 QY 299 AACACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGAGAAAGATG 359
 Db 2166 CAAATTAACAACTTAACATTTAGTTCTGCGCAAGCTTCTTTACACTAAACCGGCTTAGATC 2225
 QY 359 TTGACCACTATCCTCGTCTCAAGAGATATCAACAGAAATCCTTTGATTTTAAACAA 418
 Db 2226 CAATGGAATACCTTAAATTAACAGAGTTACTGATGTAATAATTTAAATTTCTATTA 2285
 QY 419 AATTATTGAAGTCTATTTTGTGAAACAGCTTTTTCAGCCAGTTTACAGAAAGTGCCT 478
 Db 2286 ATGCTCTTAAACAACTTTATAGACAAACTGTTTGTCTGTTCTGCTATCGAAGTTTCGCC 2345
 QY 479 CTATTTTAAACAGGAGTTCAATTTGTTATTAAGTAATCATAAAGATTTTAAAGCAGTAGCGA 538
 Db 2346 CAGTCTTACAGGTTTAACTGGATTTAT---TAAAGATAACAACTTAGCGCAGTTGCA 2402
 QY 539 CTGACTCTCATGCTATGAGCCAACTGTTTAACTCACTTTTGA---CAATACTTTCAGCAGATT 595
 Db 2403 CCGATAGTCACTGCTAGCATTAGCGTAATTCCTTGAACAAATATTTGACGAGAT 2462
 QY 596 TGATGTTAGTCTTCCAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATA 655
 Db 2463 ACAACATTTGTTATTCAGGAAAGCTTAGCTGAACATAAATAAATTTTACAGCAGTCA 2522
 QY 656 TTGAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATTT 715
 Db 2523 GTGAATCCATTTGAATGACACTCGCAATTAATCAATCTTATTTAAATTTAAAGATTTAT 2582


```

QY 716 CTTTATATACAGCGCTCTTAGAAGAAATATCCGATACAGACCGTTTATTAATGACAG 775
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2593 TATTTTATCTCGTTTACTGGAAGTAGTTACCCAGATACATCAGCTTAATTCCAACTG 2642
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 776 AATTGAGAGGAGGTGTTTCAATACCAATCCCTTCGCCACGCTATGGAACGTGCT 835
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2643 ATACTAATCAGACTTGTCATTAATTCAGAGCAATTTTACAGCGATTGACCGTGCT 2702
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 836 TCTTGATTCTTAATCTACTCAAAATGGTACTGTTAAAGCTTGAGATTACTCAAAAT---C 892
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2703 CCTACTTCTCGCGAAACCGTAATAACGTTTATTAATAATGACCCCTTGAAGATGGAC 2762
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 893 ATATTTAGCTCATCTTAATCACTCAGCTAGGTTGGTAAGTAAACAGAGATTAGATATG 952
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2763 AAGTTGAAGTATCTCTCACTCTCCTGGAAGTTGGGAATGTTTCGGAAAATGCTTTAGTC 2822
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 953 TTAGTCAGTCTGAGTAGTATTAATCACTATCAGCTTCAATCCAACTTACCTTATGAGTCT 1012
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2823 AAAGCTTTACTGCTGAAGAAATTAATAATATCTTTTAAACGTTAAATACATGATGAT 2882
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1013 TAAAGCTATTAAAGAGTGAACAGTAAATAATTCATTTTATCACCAGTTGCAACATTC 1072
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2883 TCGCGCATTTGAAGCGGATGATATTCAAATCTCTCTCTCGGCACAAATGAGACCTTTCG 2942
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1073 CCTAACACACCGCGATG 1090
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2943 TACTTCGACCAAAAGATG 2960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 24

```

US-10-282-122A-24149
; Sequence 24149, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 24149
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Listeria monocytogenes
US-10-282-122A-24149

Query Match 16.4%; Score 185.6; DB 16; Length 1146;
Best Local Similarity 50.6%; Pred. No. 1.1e-28;
Matches 555; Conservative 0; Mismatches 529; Indels 12; Gaps 4;

QY 4 ATTCATTTTCAATTTAATCGCACATTTATTTATTCATGCTTTTAAATACAACTAAACGTGCT 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 ATGAATTTTGTATTTAGCGGTGATCGCTTTGTCGAAGCAGTCATTAAGTTACTGTGCG 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 ATTAGCACTAAATAATGCCATTTCTTATTTCTTTCATCAATATAAATAATGAAAGTCACTTCT 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ATCTCTCAAGAAACAACGATTTCCAATTTTAAACGGGATATAAATAAGTCGTAATGATGAA 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 124 GGAGTAACTTTAAACAGGCTCTAAACGGTCAATATCAATTAATGAAACACATTTCTCTGTA 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GGTGTAAACACTAACTGTTAGTATTCGATATTTCCATCGAAGCATTTTATTCATTAAT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 184 ---AATGAAAATGCTGTTTCTAATTAACCTCTCCAGGAGCTATTTTATTTAGAAGTAGT 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GAAATGATGAAGTAATTTGTAAGTAGAGAGTTTTTGGTGGATTTGTTACTTCAATCAAAA 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TTTTATTAATAATTTTCAAGTTTGGCAGATATTTAGTATAAATGTTTAAAGAAATTTGAA 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TACTTTGGGATATTTGTCGTCGTTTACCAGAAAGAAATGTAGAAATTTGAAGTAACGTCT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 CAACACCAAGTTGTTTAAACAGTGTAAATCAGAGATTACCTTTAAAGGAAAGATGTT 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 AATTACCAACCAACATTTAGTTCTGTCGTCGTCCTTTTACACTAAACGGCTTTAGATCCA 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GACAGTATCTCGTCTACAGAAAGTATCAACAGAAATCTCTTTCATTTTAAACAAACAAA 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 ATGGAAATATCTTAAATTAACCTGAAGTAACAGATGGAAGAAACAAATTAATTTCAATTA 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 TTATTTGAAGTCTATTTATTTGCTGAAACAGCTTTTTCAGCCAGTTTACAAAGAAAGTCTCT 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GTACTTAAATAATTTGTTAGACAACTGTTTTTGTCTGCTTATTTAGTAACTTCTGCTCA 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ATTTTAAACAGAGTTTATTTGTTATTAAGTAAATCAATAAGATTTTAAAGCAGTAGGACT 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 GTACTTACTGTTGTAACCTGGAATTTATCAAGAAATAA---CTAAGCGCAGTTGCAACC 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GACTCTCATCTGATGAGCCAAAGTTTAACTTCTTTCGAGCA---ATACCTTACAGAGATTG 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 GATAGTCATCTGCTTAGCTTTAGTGAAATTCACCTTGAAACAGACATTTGATGAAGATAC 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 ATGCTAGTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATTT 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 AATATTTGTTTCTCTGGAAGAAAGTTTATCTGAATTTAAATAAACTTTTAGATGACGCAAGC 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 GAGACCGTTGAGGTATTTTCTCAGCAAGCCAAATCTTTGTTGAGAGTGAACACATTTCT 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 GAATCTATTGAAATGACCCCTTTCGCAACCAACCAATTTCTTTTAAATTTAAAGATTTATTA 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 TTTTATACACGCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTTAATGACAGAA 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 TTTTACTCTGTTTACTTGAAGGTAGTTTACCCAGATATATCTCGATTAATTTCAACTGAT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 TTTGAGACGGAGTTGTTTTCATACCCAAATCCCTTCGCCAGCTTATGGAACGTGCTTTC 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 ACTAAATCAGAAATTAGTCAATTAATTTCCAAAGCATTTTACAAAGCAATTTGACCGTGC 837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTTAAAGCTTGAGATTACTCAAAAT---CAT 894
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 838 CTACTTGTCTCGCAAAATCGTAATTAACGTTTATTAATTAATTAAGCTGCAAAATGCCAA 897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 ATTTTCAGCTCATGTTTAACTCAGCTGAGTTGTAAGGTAAACGAGGATTTAGATATGTT 954
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 898 GTAGAAGTATCTCTCAATTTCTCCGGAAGTTGGGAATGTTTCTGAAAATGCTTTCGCCAA 957
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 955 AGTCAGTCTGGTAGTGAATTAATCACTACAGCTTCAATCCAACTTACCTTATGAGTCTTTA 1014
 Db |||||
 QY 958 AGTTTACTGGCGAAGAAATCAAAATATCTTTTAAACGGTAATACATGATGCGCTTA 1017
 Db |||||
 QY 1015 AAAGCTATTAAAGTGAACAGTAATAAATTCATTTATACCAAGTTCGACCAATTCACC 1074
 Db |||||
 QY 1018 CGTGCTTTTGAAGGTGATGATATTCAAAATTTCTCTCAGGTACTATGAGACCAATTTGA 1077
 Db |||||
 QY 1075 CTAACACCGGCGATG 1090
 Db |||||
 QY 1078 CTTGACCAAAAGATG 1093
 Db |||||

RESULT 25

US-10-282-122A-9444
 ; Sequence 9444, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangshu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9444
 ; LENGTH: 1137
 ; TYPE: DNA
 ; ORGANISM: Bacillus anthracis
 US-10-282-122A-9444

Query Match 15.7%; Score 178.4; DB 16; Length 1137;
 Best Local Similarity 49.7%; Pred. No. 3.5e-27;
 Matches 567; Conservative 0; Mismatches 561; Indels 12; Gaps 4;
 QY 4 ATTCAATTTTCAATTAATCGACATTAATTTATTCATGCTTTAAATACAACTAAACGCTGCT 63
 Db |||||
 QY 1 ATGGCTTTTCAATTCAAAAGACATCTCTGTAAGAAGTGTACAGATGTAATGAAGGCT 60
 Db |||||
 QY 64 ATTAGCACTAAATAATGCCATTCCTATTCTTTTCATCAATAAAAAATTTGAAGTCACCTTCTACA 123
 Db |||||

Db |||||
 QY 61 GTTCTTTTCGTACAACTATTCGATCTCTTACAGAAATTAAGTTGTTCGCTACGGAAGAA 120
 QY 124 GGAGTAATCTTTAAACAGGCTCTAACGGTCAAAATATCAATTTGAAAAACATCTATTCTCT--A 180
 Db |||||
 QY 121 GGAGTTACATTAACAGGAAGCGATGCTGATATTTCGATTGAATCTTTTATCCAGTTGAA 180
 QY 181 AGTAATGAATGCTGGTTTCTGCTAATTTACCTCTCCAGGAGCTATTTTATTAGAAGCTAGT 240
 Db |||||
 QY 181 GAGGATGGAAAAGAAATTTGTAGAAGTAAACAATCAGGAAGTATTTGTTTACAGGCTAAA 240
 QY 241 TTTTATTAAATATTATTTCAGGTTTGGCAGATATAGTATAAATGTTTAAAGAAATTTGAA 300
 Db |||||
 QY 241 TATTTAGTGAATTTGTAATAAATTTGCCGAAAGAACTGTAGAAATTTCTGTCGAAAT 300
 QY 301 CAACACCAAGTTGTTTAAACAGTGGTAAATCAGAGATTTACCTTAAAGGAAAAAGATGTT 360
 Db |||||
 QY 301 CATTTAATGACAAAAATAACTTCTGGGAAATCAGAAATTTAATTTTAAATGGTTAGATTCT 360
 QY 361 GACAGTATCTCTGCTACAGAAAGTATCAACAGAAATCTCTTGTATTTTAAACAAAAA 420
 Db |||||
 QY 361 GCAGATATCCATTTGTTTACCAAAATTTGAAGAAATCATGTTTAAAGATTTCCACAGAT 420
 QY 421 TTATTGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAAGAAAGTCTCCT 480
 Db |||||
 QY 421 TTACTTAAACATATGATCAGACAACTGTTTTCAGCTCTCCACTTCTGAAACAAAGACCA 480
 QY 481 ATTTTAAACAGAGTTCA---TATTGTAATTAAGTAAATCATAAAGATTTTAAAGCAGTAGCG 537
 Db |||||
 QY 481 ATCTTGACGGTGTAAACTCTGGAAGGTATATAACAGCGAACTAACTTGTATTGCTACAGAT 540
 QY 538 ACTGACTCTCATCGTATGAGCCAGCTTTAATCACTTTTGACAACTACTTACGAGATTTG 597
 Db |||||
 QY 541 AGTCACAGTTAGCTCTTCGAAAGCAAAATTTGAGGTAATTTGAGATGAATTTCCAGGCA 600
 QY 598 ATGGTAGTTCTTCCAAGTAAATCTTTGAGAGAAATTTTCAGCAGTATTTTACAGATGATTT 657
 Db |||||
 QY 601 AATGTTGTTATTCGGGAAAAAGCTTAAATGAATTAAGCAAAATTTCTAGATGATCTGAA 660
 QY 658 GAGACCGTTGAGGTATTTTCTCACCAGGCAAAATCTTGTTCAGAGTGAACACATTTCT 717
 Db |||||
 QY 661 GAAATGTTAGATATCGTTATTACGGAGTATCAAGTATTTATTCGCTACAAAACATTTATTA 720
 QY 718 TTTTATACACGCTCTTTAGAAAGAAATTTATCCGATACAGACCGTTTATTAAATGACAGAA 777
 Db |||||
 QY 721 TTCTTCTCAAGATTTGTAGAAAGAAATTTCTGTATACAACTCAATTAATTCCTGCGAG 780
 QY 778 TTTGAGACGGAGGTGTTTTCATACCCAAATCCCTTTGCCACGCTATGGAACGTCCTTC 837
 Db |||||
 QY 781 AGTAAAAACAGATATTTTGTAAATACAAAAGAAATTTTACAAGCAATTTGATCGTCATCT 840
 QY 838 TTGATTTCTAATGCTACTCAAAATGGTACTGTTTAAAGCTTGAGATTACTCAAAATCATATT 897
 Db |||||
 QY 841 CTGTTAGTAGAGATGCTGTAATAATTTGTTGCAAAATTTATCAACTTTTAGACGAGCAATG 900
 QY 898 TCAG---CTCATGTTAACTCACCTGAGTTGGTAAAGTAAACGAGATTTTAGATATTGTT 954
 Db |||||
 QY 901 CTAGAAATTTCTCGAATTCACAGAAATCGGGAAGTAGTAGAAGAGTTCAATGTGAA 960
 QY 955 AGTCAGTCTGTTAGTGAATTTAATCACTAGCTTCAATCCAACTTACCTTATTGAGTCTTTA 1014
 Db |||||
 QY 961 AAAAGTAGATGGAGAAGAGTTAAAAATATCTTTTAGTGCAAAATATATGATGATGCACTA 1020
 QY 1015 AAAGCTATTAAAGTGAACAGTAAATAATTCATTTCTTATCACCACTTCGACCAATTCACC 1074
 Db |||||
 QY 1021 AAGCAATTAGATGATGATGAAATTAAGATTAGCTTTTACTGGAGCAATGAGACCAATT--C 1077
 QY 1075 CTAACACCGGCGATGAGGAAGAAAGTTTTATCCAAATTAATTAACACAGTACGAAACAAAC 1134
 Db |||||
 QY 1078 TTAATTCGTCGCGTAAATGATGAATCCATTATTCAATTAATTTTACCGGTTCTGACTTAC 1137
 Db |||||

```
US-09-974-300-1554
; Sequence 1554, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1554
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1554

Query Match          9.3%; Score 105.8; DB 9; Length 1072;
Best Local Similarity 47.3%; Pred. No. 4.8e-12; Indels 12; Gaps 4;
Matches 456; Conservative 0; Mismatches 497;

QY 61 GCTATTAGCACTAAATAATGCCATTCCTATTCTTCATCAATAAAATTTGAAGTCACATTCT 120
DB 58 GCGGTTCTTCAAGAACGACGATTCGGATCTTAACCGGTATTAAATCGTGGCCTCTGAT 117
QY 121 ACAGAGTAACCTTTAAACAGGGTCTAACCGGTCAAATATCAATATGAATAAAACACTATTCTCTGTA 180
DB 118 GAAAGGGTCTCTCTGACAGCGGAGCGATTCGGATATTCGATTTGAATCGCTTTATCCCGAAA 177
QY 181 AGTAATGAAT--GCTGGTTGCTAATTACCTCTCCAGGAGCTATTTTATTAGAGCT 237
DB 178 GAAGACGGCGATTTAGAGATCGTGAATTTGAACAGCCCGGAGCATTTGCTTCAAGCC 237
QY 238 AGTTTTTTTAAATATATTATTTCAAGTTTGGCAGATATTAGTATATAATGTTTAAAGAAAT 297
DB 238 CGTTTTTTCAGTGAATTTGTCAAAGCTGCCGATGTCAACAGTGAATTCAGGTTCAA 297
QY 298 GAAACAACCAAGTTGTTTAAACAGTGGTAAATCAGAGATACCTTTAAAGGAAAGAT 357
DB 298 AATCAATACTTAAACGATCATCGCTCCGCGAAAGCAGAGTTTAACTTAAACGGTTTGGAT 357
QY 358 GTTGACCAAGTATCCCTCGCTACAGAGATATCAACAGAAATCCTTTGATTTTAAACA 417
DB 358 GCAAGCGAATATCCGCTTTTGGCGCAATTTGAAGAGCATCACGCTTTTCAAATTCGACC 417
QY 418 AAATTTATGAAGTCTATTATTGCTGAAACAGCTTTTGCAGCCAGTTTACAAAGAAAGTCGT 477
DB 418 GATCTGCTGAAACAACTGATCGCCAAACCGTTTGTGAGTGTCCACCTCAGAACACGC 477
QY 478 CCTATTTAAACAGGAGTTCATATTGTTATAGTAATCAATAAGATTTTAAAGCAGTAGCG 537
DB 478 CCAATCTTCAACAGGTGTAATGGAATGTCACTGGCGGT---GAATTAATATGCACTGCA 534
QY 538 ACTGACTCTCATCGTATAGGCCAAAGTTTATCACTTTGGACAACTTTCAGCAGAT--- 594
DB 535 ACGGATATGTCATCGCTTTCGCGTAAAGGAAAGCTAAGCTCGACATTTAAACGAAGACAGTTCA 594
QY 595 TTGATGGTAGTCTTCTTCAAGTAAATCTTTTGAGAGAAATTTTACGAGTATTTACAGATGAT 654
DB 595 TACATGTGCTCATCCCGAGAAAAGCTTTACCGAGTTCAGCAAAATCTTGTATGACCAT 654
QY 655 ATTGAGACGGTTGAGGTATTTTCTCAACAGCCAAATCTTTGTCAGAGTGAACACAT 714
DB 655 CAGGAGCTTGTGATATTGTGATTACCGAAACACAAAGTGTCTTTTAAACAAACAAACGTT 714
QY 715 TCTTTTATACCGCTCTTAGAAGGAAATTTATCCGATACAGACCGCTTTTATTATGACA 774
```

Db 715 CTGTTTTTCTCCAGGCTTTTGACGGAACATATCCGGATACGAACCGCTGATTCCTCAG 774

QY 775 GAATTTGAGACGAGGTTGTTTTCAATACCAATCCCTTCCGACGCTATGGAACGTGCC 834

Db 775 GAAAGCAAAACGAACTTGATTGTCAATACTAAGGAATTTTCTCCAGGCGATCGACAGGCT 834

QY 835 TTCTTGATTTCTAATGCTACTCAAAATGGTACTGTTAAGCTTGAGATTACTCAAAATCA- 893

Db 835 TCGCTTTTGGCGAGAGGAGCAACAAACGTTGGTGAACATTTCCGACGCCCAATGAG 894

QY 894 --TATTTTCAGTCAATGTTTAACTCACTGAGGTTCGTAAAGTAAACGAGGATTTAGATATT 951

Db 895 TCGATCGAAATTTCTTCAAACTCTCCGAAATCGAAAGGTTGTTGAACCGTGAATGCC 954

QY 952 GTTAGTCAGTCTGCTAGTGAATTTAACTATCACTGCTTCAATCAACTTACCTTATTGAGTCT 1011

Db 955 GAGCAAAATCGAAGGGGAGGACTTAAAGATATCTTTAGTCCGAAATATGTGCGGAACCA 1014

QY 1012 TTAATA 1016

Db 1015 TAAAA 1019

RESULT 27

US-10-282-122A-17164

; Sequence 17164, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA 034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 17164

; LENGTH: 1206

; TYPE: DNA

; ORGANISM: Clostridium difficile

US-10-282-122A-17164

Query Match	9.1%;	Score 103.2;	DB 16;	Length 1206;
Best Local Similarity	47.6%;	Prod. No. 1.7e-11;		
Matches 436;	Conservative 0;	Mismatches 468;	Indels 12;	Gaps 4;
QY	216	AGGAGCTATTTTATTAGAACGTAGTGTTTTATTAAATATTATTCAAGTTTGCCAGATAT	275	
Db	294	AGGTGATGTTGTTGTAGATGCAAGCTATTTGGAGATATAATTAGAAAATTACCAATTC	353	
QY	276	TAGTATAAATGTTTAAAGAAATTGAACAACACCAAGTTGTTTAAACCAAGTGGTAAATCAGA	335	
Db	354	CTTTGTTTGAGATAGAAACTGATTTCAGAAAAATAATTACATAAATTTGTGTAAATTC	413	
QY	336	GATTACCTTTAAAGGAAAGATGTTGCCAGTATCCTCGTCTCAAGAAGATATCAACAGA	395	
Db	414	ATTTAAAATTAAGGTTATGCGCAAGAAGATTTTCCTAAATTCACGAATTAATGAAGA	473	
QY	396	AAATCCTTTGATTTTAAAAAATAAATTAATTGAAGTCTATTATTGCTGAAACAGCTTTTGC	455	
Db	474	AGATTTTATAGCATACCTCAAGAAATCCTAAAAAATATGATTAACAACAACTGTATTTCG	533	
QY	456	AGCCAGTTTACAAGAAAGTCTGCTCTATTTTAAACAGGAGTTCATATTCGTTTAACTATCA	515	
Db	534	TATCTCACAGCACCAAAACCTGTTTAAATCGGAG--AACTTTTGAAGAAATTTGTCGGT	590	
QY	516	TAAAGATTTTAAAGCAGTAGCGACTCTCATCGTATGAGCCAAAGCTTTTAATCACTTT	575	
Db	591	TAGAAACTTTAAATTTAGTAGCAATTTGATGGATATAGATTAGCTGTGTAATAAGTTTCTCGT	650	
QY	576	GGACAACTTTCACGAGATTTGATGGTAGTTCCTCCAAGTAATATCTTTGAGAGATTTTC	635	
Db	651	TGATGTTTAACTGAAAATTAAGAATCAATTTCTCGGAAGACACTTATAGATGTAAA	710	
QY	636	AGCAGTATTACAGATGATATTGAGACCGTTGAGGTTATTTTCTCAACCAAGCCAAATCTT	695	
Db	711	TAGTTTCTTTCTCGAGA--GGATAATGTTAAAGTTGGATTTAAATGAAAAAATGCTAT	767	
QY	696	GTTCCAGAGTGAACACACTTTCTTTTATACAGCCTCTTAGAGAGGAATTTATCCGATAC	755	
Db	768	TTTTTAAATTAATGATACGAAATTTATACAGATTCGTTGAAGGAGATTTTATTGATTA	827	
QY	756	AGCCGTTTATTAATGACAGAAATTGAGCGGAGGTTGTTTTCAAATACCAATCCCTTCG	815	
Db	828	TAAAAAATTTTGCACAGAGAACATAACAGCAGAGTTAAATTAATTAATCAAGAACCTTTT	887	
QY	816	CCAGCTATGGAACGTCCTTCTGATTTCTTAATGCTACTCAAAATGGTACTGTTAAGCT	875	
Db	888	AAATGATATAGAGAGCGTCTTTATGTCTCAATCAGAAAAAATAATCTTATAAAGTT	947	
QY	876	TGAGATTACTCAAAATCATATTTTCAGCTCATGTTAACTCACTCAGGTTTGGTAAAGTAAA	935	
Db	948	ATCTATAAGAGATAAAGTAATGCGCAATTACTTCTAATACA--GAAAAAGGAATGTGTA	1004	
QY	936	CGAGGATTTAGATATTTGTTAGTCAGTCTGGTAGTGATTAACTATCACTGCTTCAATCCAAC	995	
Db	1005	TGAAGAAGTTTGAATATGATTTAGA--TGAGATTTACCTAGATATAGCTTTCAATTTCTAG	1061	
QY	996	TTACCTTTATTGAGTCTTTTAAAGCTATTTAAAGTGAAACAGTAAAAAATTCATTTCTTATC	1055	
Db	1062	ATATTTTATCGAAGGTTTGAAAAATAATAGATAATGAGGAATATTTATAGAAATTTACTAC	1121	
QY	1056	ACCAGTTTCGACCAATTCACCTTACACACGCGGATGAGGAAGAAAGTTTTATCCAAATTAAT	1115	
Db	1122	AAATGTAATTCCTTTGTTAAATTTAAACCAACAGATGATGTTTAAATATATTACTTACTGCT	1181	
QY	1116	TACACCAGTACGAACA	1131	
Db	1182	TCCAGTAAGAATATCA	1197	

RESULT: 28
US-10-282-122A-16193
; Sequence 16193, Application US/10282122A
; Publication No. US20040029129A1

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsuo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16193
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; US-10-282-122A-16193

```

Query Match	9.0%; Score 102.4; DB 16; Length 1125;
Best Local Similarity	48.1%; Pred. No. 2.5e-11;
Matches	416; Conservative 0; Mismatches 436; Indels 12; Gaps 4;
Qy	217 GGAGCTATTTTATTAGNAGCTAGTCTTTTTTATTAAATATATTTTCAAGTTTGGCCAGATATT 276
Db	217 GGAACATATAGTTTTAGATGCTACAGCTTTTTTGGTGAGATATAAGAAAATTTACCTAATGAT 276
Qy	277 AGTATAAACTGTTAAAGAAAATTGAACAACCAAGTTGTTTTTAAACCAGTGCTGAATACAGAG 336
Db	277 CTAAAGAAATAATACGTTAGAGAGATAATCTATAGAAATATATGTCAAAATTTCTAGA 336
Qy	337 ATTAACCTTAAAGGAAAAGATGTTGACCAAGTATCCTCGTCTACAAGAAGTATCAACAGAA 396
Db	337 TTTTAACTTAATATATATGATGAATCCAGGTGAGTTTTCTTAATCCAGCTATATCAATGAAAAT 396
Qy	397 AATCCCTTTGATTTTAAAAACAAAATTTATTGAAGTCTATTATTGCTGNAACAGCTTTTGCA 456
Db	397 ATGATATTTTCTATAGGTGAAGTAAATTAATAAAATATGATTAAGAGAACTATATTGTCT 456
Qy	457 GCCAGTTTACAAGAAAGTCGTCTCTATTTTAAACAGGAGTTCATATGTGTTATTAAAGTAATCAT 516
Db	457 ACAGCTCAAGATGAACAAGACCCTATTCTTACAGGTGTTTTATTTCAAATA---AAGAC 513
Qy	517 AAGATTTTAAAGCAGTAGGCACTGACTCTCATCGTATGAGCCACGTTTAAATCACTTTG 576
Db	514 AAAATGTTAAATTTAGTAGCTTTTAGATGGCTATAGGTTAGCTTTTAAAG---AAGCGAAGTT 570

QY 577 GACAATCTTACAGAGATTGAGTGTCTTCCAGTAATAATCTTTGAGAGAAATTTTCA 636
DB 571 GTTGATAATGATAATACGATAAATCTGTATTTCAGGGGAAACATTAATGAAGTTTCA 630
QY 637 CGAGTATTACAGATGATATTGAGACCGTTGAGGTATTTTCTCACCAGCCAAATCTTG 696
DB 631 AAAATTTTAGAAGAAGAAATGAAATGTTAATATACATTTTACTCTTAAATCATATTTTA 690
QY 697 TTCAGAAGTGAAACACATTTCTTTTATACACGCCCTCTTAGAGGAAATTTATCCGATACA 756
DB 691 TTTAGTATAGGAGAAACAAAATAATTTCAAGATTATTAGAGGAGAAATTTATTAGCTAT 750
QY 757 GACCGTTTATTAATGACAGAAATTTGAGACGGAGGTGTTTTCATACCAATCCCTCCG 816
DB 751 AAGCTATAATAACCTGAAGAATTTAACTTAAATAAATAGCTTAAAGATCAGAGCTTTTA 810
QY 817 CACGCTTATGGAACGCTCTCTTGTATTCTTAATGCTACTCAAAATGGTACTGTTAAGCTT 876
DB 811 AATCTATAGAAAGACCTCTCTTATGGCTAAGAGGAATACTAAATTTAGTTAAATTT 870
QY 877 GAGATTACTCAAAATCATATTTTCACTCATGTAACTCACCTGAGGTGGTAAGTAAAC 936
DB 871 GATTTCTCAGATGATAAATTTGTAATAACATCAAAATCTC---AATTAGGAATGTTTGA 927
QY 937 GAGGATTTAGATATTGTTAGTCAGTCTGTAGTGTGATTTTAACTATCAGCTTCAATCCA 996
DB 928 GAAGAATTTAAAGTTGTATTGCA---GGAGAAGATTTTCAAAATTTGCAATTTAAATTC 984
QY 997 TACCTTATGAGTCTTTAAAGCTATTAAAGTGAAACAGTAAATAATTCATTTCTTATCA 1056
DB 985 TATCTTTTAGATGTTGAAACTATGGAAGATAATGAGGTGTATTAGAAATTTTCAAGT 1044
QY 1057 CCAAGTTCGACCAATCCACCTTACA 1080
DB 1045 AGTGAAGTCTTGTATTATAAA 1068

RESULT 29

US-10-671-403-173
; Sequence 173, Application US/10671403
; Publication No. US20040038289A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalimi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/671,403
; PRIOR APPLICATION NUMBER: 2003-09-25
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 173
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-10-671-403-173

Query Match 7.8%; Score 88; DB 16; Length 992;
Best Local Similarity 46.5%; Pred. No. 2.4e-08;

Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;
QY 214 CCAGAGACTATTTTATTAGAACTAGTTTTTTTATTATTAATTTTCAAGTTTGCAGAT 273
DB 73 CCGGGGAGCATCGTACTCGACGCGCTTTTCTCTGAAATCGTGAAAAAATCGCGCAA 132
QY 274 ATTAGTATAATGTTTAAAGAAATTTGAACAAACCAAGTTGTTTTAACAGTGGTAAATCA 333
DB 133 CAAACGGTGGAAATCGAAACCGAAGACAACTTTTGTACGATCATCCGCTCGGGGCACTCA 192
QY 334 GAGATTACTTTAAAGAGAAAGATGTTGACCACTATCTCGTCTACAAAGAACTATCAACA 393
DB 193 GAATTCGCCCTCAATGGCTAAACGCCGACGAAATATCCGCGCTCGCGCAATTTGAAGNA 252
QY 394 GAAATCTTTGATTTTAAACAAATTTATGAAGTCTATTTATTTGCTGAAACAGCTTTT 453
DB 253 GAAACGCTGTTTCAAAATCCCGCTGATTTATTGAAACCGTGTATTCGGCAACCGTGTTC 312
QY 454 GCAGCCAGTTTACAAGAAAGTCTGCTATTTTACAGGAGTTTCAATTTGTATTAAAGTAAT 513
DB 313 GCCGTTTCTACATCGGAAACCGGCCCAATCTTGACAGGTGTCAACTGGAAAGT---TGAA 369
QY 514 CATAAAGATTTTAAAGCAGTAGCAGTCTCTCATCTATGATGAGCAACGTTTAAATCACT 573
DB 370 CATGGCAGCTTGTGTCACAGCGCAAGTCTCTCTCTTTCCCGCTGCTTGAACGCAACTATCG 429
QY 574 TTGGACA---ATACCTTTCAGCAGATTTGATGAGTGTCTTCCAAAGTAAATCTTTGAGAGAA 630
DB 430 ATTGAGTCGGAAATGAAGTATATACAAACGTCGTCTCTCTTCCCGCTGCAATCTTAATGAG 489
QY 631 TTTTCAGCAGTATTTACAGATGATATTGAGACCGTGTAGGTATTTTCTCACCAGCCAA 690
DB 490 CTCAGCAAAATTTTGGATGACGGCAACCAACCGGTGGACATCGTCATGACAGCCAACTCAA 549
QY 691 ATCTGTTTTCAGAAAGTGAACACATTTCTTTTATACACGCTCTTATAGAGGAAATTTATCCC 750
DB 550 GTGCTATTTAAGCGCAGACCTTCTCTCTTTCCCGCTGCTTGAACGCAACTATCG 609
QY 751 GATACAGACCGTTTATTAATGACAGATTTGAGACGGAGGTGTTTTCAATACCAATCC 810
DB 610 GAGACGGCCGCTTGATTTCCAAACAGAAACGACCATGATCGTCAATGCAAAAGAG 669
QY 811 CTTGCCACGCTATGGAACGCTCTCTTCTGATT---TCTAATGCTACTCAATGCTACTG 868
DB 670 TTTCTGACGCAATTCGACCGGCTCTTCTGCTGCTGAGAGGAAAGCAACCTGTTG 729
QY 869 TTAAGCTTTGAGATTACTCAAAATCATATTTTTCAGCTCATGTTAACTCACTCAGGTGTTA 928
DB 730 AAACGTGACGACGCTTCTGGAGGAATGCTCGAAATTTCTCGATTTCTCCGAGATCGGA 789
QY 929 AGGTAACGAGGATTTAGATATTGTTAGTCACTGCTGTTAGTATTAATCACTATCAGTTCA 988
DB 790 AAGTGACGGAGCAGCTGCAAAACGGAGTCTCTTGAAGGGGGAAGAGTTGAACATTTCTGTTCA 849
QY 989 ATCCAACCTTACTTATGAGTCTTTTAAAGCTATTTAAAGTGAACAGACAGTAAATTTCA 1046
DB 850 CGCGAAATATATGATGACGCGTTGCGGGCGCTTGTATGGAACAGACATTTTCAAAATCA 907

RESULT 30

US-10-671-419-173
; Sequence 173, Application US/10671419
; Publication No. US20040038290A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalimi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/671,403
; PRIOR APPLICATION NUMBER: 2003-09-25
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 173
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-10-671-403-173

; FILE REFERENCE: 22221/1030
 ; CURRENT APPLICATION NUMBER: US/10/671,419
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US/09/716,964A
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416
 ; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 173
 ; LENGTH: 992
 ; TYPE: DNA
 ; ORGANISM: Bacillus stearothermophilus
 US-10-671-419-173

 Query Match 7.8%; Score 88; DB 16; Length 992;
 Best Local Similarity 46.5%; Pred. No. 2.4e-08;
 Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;

 QY 214 CCAGGAGCTATTTTATTAGAACTAGTTTTTTTATTATATATTTTCAAGTTTGCAGAT 273
 DB 73 CCGGGAGCATCGTACTCAGCGCGCTTTTTCTCTGAAATCGTGAAAAAAGTCCGCA 132

 QY 274 ATTAGTATAAATGTTTAAAGAAATGAAACCAACCAAGTTGTTTAAACAGTGGTAAATCA 333
 DB 133 CAAACGGTGGAAATCGAAACCGAAGACAACTTTTTCAGATCATCCGCTCGGGCAGTCA 192

 QY 334 GAGATTACCTTAAAGGAAAGAGTGTGACAGTATCTCTGCTCTACAAGATATCAACA 393
 DB 193 GAATTCGGCTCAATGGCTTAACCGCCAGCAATATCCGCGCTCCGCAATTTGAAGA 252

 QY 394 GAAATTCCTTTGATTTTAAAAAAGTGTGAGTGTGTTTAAAGTCTATTATGCTGAAACAGCTTTT 453
 DB 253 GAAACAGTGTGTTCAAAATCCCGCTGATTTATTGAAACCGTGAATTCGGCAACCGTGTTC 312

 QY 454 GCAGCCAGTTTACAGAAAGTTCGTCTATTTTAAACAGGAGTTCATATTGTTAATGAAT 513
 DB 313 GCGGTTTCTACATCGAAACCGCCCAATCTTGACAGTGTCACTGGAAGT---TGAA 369

 QY 514 CATAAAGATTTTAAAGCAGTAGCAGTCTCATCTGATGACCAACCGTTTAAATCACT 573
 DB 370 CATGGCAGCTTGTCTGCACAGCGACCGACAGTATCGCTTAGCCATGCGCAAAAGTGA 429

 QY 574 TTGGACA---ATACCTTCAGCAGATTGATGAGTGTCTTCCAGTAAATCTTTTGAGAA 630
 DB 430 ATTGAGTCGGAATAAGATATCATACAACTGCTGATCCCTCGAAAGTCTTAATGAG 489

 QY 631 TTTTCAGCAGTATTTACAGATGATTTGAGACCGTGTAGGATTTTTTCTACCAAGCAA 690
 DB 490 CTCAGCAAAATTTTGGATGACGGCAACCAACCGGTGGACATCGTCAATGACAGCAATCA 549

 QY 691 ATCTGTGTTTCAAGTGAACATTTCTTTTATACAGCTCTTTAGAGGAATATATCC 750
 DB 550 GTGCTATTTAAGCGGAGCACTTCTCTTTTCCCGCTGCTTGAACGCAATATPCG 609

 QY 751 GATACAGACCGTTTATTAATGACAGAAATTTGAGACGGAGGTGTTTTCAATACCAATCC 810
 DB 610 GAGACGGCCCGCTTGATTTCCAAACAGAAAGCAAAACGACCATGATCGTCAATGCAAAAGAG 669

 QY 811 CTTGCCACGCTATGGAACGCTCTCTTGATT---TCATGCTACTCAAAATGGTACTG 868
 DB 670 TTTCTGAGGCAATCGACCGAGCTCTTCTGCTGCTGAGAGGAAGAAACAGTTGTG 729

 QY 869 TTAAGCTTGAGTACTCAAAATCATATTTTCAAGTCTCATGTTAACTCACCTGAGTTGGTA 928
 DB 730 AAATCTGACGAGCTTCTCGAGGAATGCTCGAAATTTCTTCGATTTCTCGAGATCGGA 789

 QY 929 AGGTAAACGAGGATTTAGATTTGTTAGTCTGAGTCTGATGATTTAACTATCAGTTCA 988

Db 490 CTCAGCAAAATTTTGGATGACGGCAACCCCGGTGGACATCGTCATGACAGCCCAATCAA 549
 Qy 691 ATCTGTTTCAAGTGAACACATTTCTTTTATACAGCCTCTTAGAAGGAAATATCC 750
 Db 550 GTGCTATTAAAGCCGAGCACCTTCTCTTTTCCCGCTGCTTGACGGCAACTATCG 609
 Qy 751 GATCAGACCGTTTATTAATGACAGAAATTTGAGCGGAGGTTGTTTCAATACCAATCC 810
 Db 610 GAGACGGCCCGTTGATTTCACAGAAAGCAACGACCATGATCGTCAATGCAAAAGAG 669
 Qy 811 CTTCGCCAGCTATGAAAGTCCCTCTTCTGATT--TCTAATGCTACTCAAAATGCTACTG 868
 Db 670 TTTCTGAGCAATCGACGCGCTCTTCTGCTGCTGAGAGGAAAGCAACGCTG 729
 Qy 869 TTAAGCTTTGAGATTACTCAAAATCATATTTTCAAGCTCATGTTAACTCACTGAGTTGGTA 928
 Db 730 AAATGACGACGCTTCTCGAGGAATGCTCGAAATTTCTCGATTCTCCGAGATCGGA 789
 Qy 929 AGGTTAAACGAGGATTAGATATTGTTAGTCACTCTGGTAGTGATTTAACTATCAGCTTCA 988
 Db 790 AAGTGACGAGCAGCTGCAACCGAGTCTCTTTGAAAGGGGAGAGTTGAACATTTCTGTTCA 849
 Qy 989 ATCCAACTTACCTTATGAGTCTTTAAAGCTATTAAAGCTATTAAAGTGAACAGTAAATTC 1046
 Db 850 GCGGAAATATATGATGACGCTTGGCGGCGCTTGATGGAACAGACATTTCAAATCA 907

RESULT 32
 US-10-671-134-173
 ; Sequence 173, Application US/10671134
 ; Publication No. US20040043415A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Yuzhakov, Alexander
 ; APPLICANT: Yurieva, Olga
 ; APPLICANT: Jeruzalmi, David
 ; APPLICANT: Bruck, Irina
 ; APPLICANT: Kuriyan, John
 ; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
 ; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
 ; FILE REFERENCE: 2221/1030
 ; CURRENT APPLICATION NUMBER: US/10/671,134
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US/09/716,964A
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416
 ; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 173
 ; LENGTH: 992
 ; TYPE: DNA
 ; ORGANISM: Bacillus stearothermophilus
 US-10-671-134-173

Query Match 7.8%; Score 88; DB 16; Length 992;
 Best Local Similarity 46.5%; Pred. No. 2.4e-08;
 Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;
 Qy 214 CCAGGAGCTATTTTATAGAACTAGTTTATTAATATTTTCAAGTTTGCAGAT 273
 Db 73 CCGGGAGCATCGTACTCGAGCGCGCTTTTCTCTGAAATCGTGAATACTGCGGCA 132
 Qy 274 ATTAGTATAAATGTTAAAGAAATTTGAACACCAAGTTGTTTAAACAGTGGTAAATCA 333
 Db 133 CAAACGGTGAATCGAAACGGAAGACACTTTTTCAGCATATCCGCTCGGGCACTCA 192

Qy 334 GAGATTACCTTAAAGGAAAGATGTTGACAGATATCTCTGCTTACAAGAAGTATCAACA 393
 Db 193 GAATTCGCGCTCAATGGGCTAAACCGCGCAAGATATCCGCGCTGCGCCNAAATTTGAAGAA 252
 Qy 394 GAAATCCTTTGATTATTAACAAACAAATTTATTTGAAGTCTATTTATGCTGAAACAGCTTTT 453
 Db 253 GAAACGCTGTTTCAATCCCGCTGATTTATTTGAAACCGGTGATTCGGCAACCGTGTTC 312
 Qy 454 GCAGCGAGTTTACAAGAAAGTCTCTATTTTAAACAGAGTTTCATATTTGTTATTAAGTAAT 513
 Db 313 GCGGTTTCTACATCGGAAACGCGCCCAATTTGACAGGTGTCAACTGGAAGT---TGAA 369
 Qy 514 CATAAAGATTTTAAAGCAGTAGCGACTCTCATCTGATGATGAGCAACAGTTTATCACT 573
 Db 370 CATGCGAGCTTGTCTGCACAGCGACCGACAGTCTAGCTTTAGCATCGCGCAAGTGA 429
 Qy 574 TTGACA---ATACTTTCAGCAGATTGATGTTGATGTTCTTCCAAAGTAAATCTTTTGAGAGAA 630
 Db 430 ATTGAGTCGGAATGAATGATATCATCAACGCTGTCATCCCTGGGAAAGTCTTAAATGAG 489
 Qy 631 TTTTCAGCAGTATTACAGATGATATTGAGACCGTTGAGGTTATTTTCTCACCAGCCAA 690
 Db 490 CTCAGCAAAATTTTGGATGACGGCAACCAACCGGTGGACATCGTCATGACAGCCCAATCAA 549
 Qy 691 ATCTGTTTCAAGTGAACATTTCTTTTATACAGCCTCTTAGAGAGGAAATTTATCCC 750
 Db 550 GTGCTATTAAAGCGGAGCACCTTCTCTTTTCCCGGCTGCTTGACGCGCAACTATCCG 609
 Qy 751 GATACAGACCGTTTATTAAATGACAGAAATTTGAGACGAGGTTGTTTTCATACCCCAATCC 810
 Db 610 GAGACGCGCCCTTGATTTCCAAACAGAAAGCAACGACCATGATCGTCAATGCAAAAGAG 669
 Qy 811 CTTCGCCAGCTATGGAACGCTCTTCTGATT--TCTAATGCTACTCAAAATGTTACTG 868
 Db 670 TTTCTGACAGCAATCGACGCGCTTCTGCTGCTGAGAGGAAAGCAACACCTGTG 729
 Qy 869 TTAAGCTTGAGATTACTCAAAATCATATTTTCAAGCTCATGTTAACTCACTGAGTTGGTA 928
 Db 730 AAATGACGACGCTTCTCGAGGAATGCTCGAAATTTCTCGATTCTCCGAGATCGGA 789
 Qy 929 AGGTTAAACGAGGATTAGATATTGTTAGTCACTCTGTTAGTATTAATCACTACAGTTCA 988
 Db 790 AAGTGACGAGCAGCTGCAACCGAGTCTCTTTGAGGGGAGAGTTGAACATTTCTGTTCA 849
 Qy 989 ATCCAACTTACCTTATGAGTCTTTTAAAGCTATTAAAGCTATTAAAGTGAACAGTAAATTC 1046
 Db 850 GCGGAAATATATGATGACGCTTGGCGGCGCTTGATGGAACAGACATTTCAAATCA 907

RESULT 33
 US-10-673-098-173
 ; Sequence 173, Application US/10673098
 ; Publication No. US20040048309A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Yuzhakov, Alexander
 ; APPLICANT: Yurieva, Olga
 ; APPLICANT: Jeruzalmi, David
 ; APPLICANT: Bruck, Irina
 ; APPLICANT: Kuriyan, John
 ; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
 ; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
 ; FILE REFERENCE: 2221/1030
 ; CURRENT APPLICATION NUMBER: US/10/673,098
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US/09/716,964A
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416

; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 173
 ; LENGTH: 992
 ; TYPE: DNA
 ; ORGANISM: Bacillus stearothermophilus
 US-10-673-098-173

Query Match 7.8%; Score 88; DB 16; Length 992;
 Best Local Similarity 46.5%; Pred. No. 2.4e-08;
 Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;
 Qy 214 CCAGGAGCTATTTTATTAGAGCTAGTTTATTAATATATTTTCAAGTTTGCAGAT 273
 Db |||||
 Qy 73 CCGGGAGCATCGTACTGCGGCGCTTTTCTCGAATCGTGAATACTGCCGAA 132
 Db |||||
 Qy 274 ATTAGTATAATGTTTAAAGAAATGAAACACCAAGTTGTTTAAACAGTGGTAAATCA 333
 Db |||||
 Qy 133 CAACCGTGGAAATCGAAACGGGAAGACAACTTTTACGATCATCGCTCGGGCACTCA 192
 Qy 334 GAGATTACCTTAAAGGAAAGATGTTGACCATCTCTCGTCAAGAAATGATCAACA 393
 Db |||||
 Qy 193 GAATTCGGCTCAATCGGCTAAACGGCGAGATATCCGGCTGCGGCAAAATGAAGAA 252
 Qy 394 GAAATCCCTTTGATTTTAAACCAAAATTTGAACTCTATTGCTGAAACAGCTTTT 453
 Db |||||
 Qy 253 GAAACCGTGTTCATAATCCGCTGATTTATTGAAACCGTATTCGGCAACGGTGTTC 312
 Qy 454 GCAGCCAGTTTACAGAAAGTCTCTATTTTTAAACAGGAGTTTCATATTGTTAAGTAAT 513
 Db |||||
 Qy 313 GCGGTTTCTACATCGGAAACGGGCGCAATCTTTGACAGGTGTCAACTGGAAAGT---TGAA 369
 Qy 514 CATAAAGATTTTAAAGCAGTAGCGACTCTCATCGTATCAGCAACGTTTAAATCACT 573
 Db |||||
 Qy 370 CATGGCGAGCTTGTGACACGGGACCGACAGTCTATCGTTAGCCATCGGCAAGTGAAA 429
 Qy 574 TTGGACA---ATACCTTACAGCAATTTGATGGTATTTTCCAAAGTAAATCTTTGAGAAA 630
 Db |||||
 Qy 430 ATTGAGTCGGAATAATGAAGTATCATACACGCTGTCTATCCCTGGAAAAAGTCTTAATGAG 489
 Qy 631 TTTTCAGCAGTATTTACAGATGATTTGAGACCGTTGAGGTATTTTCTCACCAGCCAA 690
 Db |||||
 Qy 490 CTCAGCAAAATTTTGGATGACGGCAACACCGGTTGGACATCGTCATGACGCCATCAA 549
 Qy 691 ATCTTGTTCAGAAAGTAAACAAATTTTATACACGCTCTTTAGAGGAAATTTATCCC 750
 Db |||||
 Qy 550 GTGCTATTTAAGCGGAGCACCTTCTCTCTTTTCCGGCTGCTTGAACGCACTATCCG 609
 Qy 751 GATACAGACCGTTTAAATGACAGAAATTTGAGACGGAGGTGTTTTCATACCCAAATCC 810
 Db |||||
 Qy 610 GAGACGGCGCTTGAATCCACAGAAAGCAACGACCATGATCGTCAATGCAAAAGAG 669
 Qy 811 CTTGCCACGCTATGGAACGTCCTTCTTTGATTT---TCTAATGCTACTCAAAATGTTACTG 868
 Db |||||
 Qy 670 TTTCTGAGGCAATCGACGGGCTCTTCTGCTCGAGAGGAGGACACAGTTGTTG 729
 Qy 869 TTAAGCTTGAATTAATAATCAATTTTCAAGTCAATTTTAACTCACTGAGTTGGTA 928
 Db |||||
 Qy 730 AAATGACGAGCTTCTCGGAGGAATGCTGAAATTTTCTCGATTTCTCCGAGATCGGA 789
 Qy 929 AGGTAAACGAGGATTTAGATATTGTTAGTCACTGCTGGTGGTATTTAACTATCAGTTCA 988
 Db |||||
 Qy 790 AAGTACGAGCAGCTGCAACGAGGCTCTTTGAAGGGGAGAGTTGAACATTTTCGTCA 849
 Qy 989 ATCCAACTTACCTTATTGAGTCTTTTAAAGCTATTTTAAAGTGAACAGTAAATTTCA 1046
 Db |||||
 Qy 850 GCGGAAATATATGATGACGCGTTGCGGCGCTTGTGGAACAGACATTTCAATCA 907

RESULT 34
 US-10-672-638-173
 ; Sequence 173, Application US/10672638

; Publication No. US20040077012A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Yuzhakov, Alexander
 ; APPLICANT: Yuzhakov, Olga
 ; APPLICANT: Jeruzalmi, David
 ; APPLICANT: Bruck, Irina
 ; APPLICANT: Kuriyan, John
 ; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
 ; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
 ; FILE REFERENCE: 22221/1030
 ; CURRENT APPLICATION NUMBER: US/10/672,638
 ; PRIOR FILING DATE: 2003-10-09
 ; PRIOR APPLICATION NUMBER: US/09/716,964A
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416
 ; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 173
 ; LENGTH: 992
 ; TYPE: DNA
 ; ORGANISM: Bacillus stearothermophilus
 US-10-672-638-173

Query Match 7.8%; Score 88; DB 16; Length 992;
 Best Local Similarity 46.5%; Pred. No. 2.4e-08;
 Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;
 Qy 214 CCAGGAGCTATTTTATTAGAGCTAGTTTATTAATATATTTTCAAGTTTGCAGAT 273
 Db |||||
 Qy 73 CCGGGAGCATCGTACTGCGGCGCTTTTCTCGAATCGTGAATACTGCCGAA 132
 Qy 274 ATTAGTATAATGTTTAAAGAAATTTGAAACACCAAGTTGTTTAAACAGTGGTAAATCA 333
 Db |||||
 Qy 133 CAACCGTGGAAATCGAAACGGGAAGACAACTTTTACGATCATCGCTCGGGCACTCA 192
 Qy 334 GAGATTACCTTAAAGGAAAGATGTTGACCATCTCTCGTCAAGAAATGATCAACA 393
 Db |||||
 Qy 193 GAATTCGGCTCAATCGGCTAAACGGCGAGATATCCGGCTGCGGCAAAATGAAGAA 252
 Qy 394 GAAATCCCTTTGATTTTAAACCAAAATTTTGAAGTCTATTGCTGAAACAGCTTTT 453
 Db |||||
 Qy 253 GAAACCGTGTTCATAATCCGCTGATTTATTGAAACCGTATTCGGCAACGGTGTTC 312
 Qy 454 GCAGCCAGTTTACAGAAAGTCTCTATTTTTAAACAGGAGTTTCATATTGTTAAGTAAT 513
 Db |||||
 Qy 313 GCGGTTTCTACATCGGAAACGGGCGCAATCTTTGACAGGTGTCAACTGGAAAGT---TGAA 369
 Qy 514 CATAAAGATTTTAAAGCAGTAGCGACTCTCATCGTATGAGCAACGTTTAAATCACT 573
 Db |||||
 Qy 370 CATGGCGAGCTTGTGACACGGGACCGACAGTCTATCGTTAGCCATCGGCAAGTGAAA 429
 Qy 574 TTGGACA---ATACCTTACAGCAATTTGATGGTATTTTCCAAAGTAAATCTTTGAGAAA 630
 Db |||||
 Qy 430 ATTGAGTCGGAATAATGAAGTATCATACACGCTGTCTATCCCTGGAAAAAGTCTTAATGAG 489
 Qy 631 TTTTCAGCAGTATTTACAGATGATTTGAGACCGTTGAGGTATTTTCTCACCAGCCAA 690
 Db |||||
 Qy 490 CTCAGCAAAATTTTGGATGACGGCAACACCGGTTGGACATCGTCATGACGCCATCAA 549
 Qy 691 ATCTTGTTCAGAAAGTAAACAAATTTTATACACGCTCTTTAGAGGAAATTTATCCC 750
 Db |||||
 Qy 550 GTGCTATTTAAGCGGAGCACCTTCTCTCTTTTCCGGCTGCTTGAACGCACTATCCG 609
 Qy 751 GATACAGACCGTTTAAATGACAGAAATTTGAGACGGAGGTGTTTTCATACCCAAATCC 810
 Db |||||
 Qy 610 GAGACGGCGCTTGAATCCACAGAAAGCAACGACCATGATCGTCAATGCAAAAGAG 669

QY 811 CTTCGCCAGGCTATGGAAGCGTCTTCTTGATT--TCTAATGCTACTCAAAATGGTACTG 868
DB 670 TTCTGCGAGCAATCGACCGAGCGCTCTTGCTGCTCGAGAAGGAAGCAACAACTTGTG 729
QY 869 TTAAGCTTGAGATTACTCAAAATCATATTTTCAGCTCATGTTAACTCACTGAGGTTGGTA 928
DB 730 AAATGACGAGCGCTTCCTGGAGGAATGCTCGAAATTTCTTGATTTCTCCGAGATCGGGA 789
QY 929 AGGTAACGAGGATTAGATATTTAGTCTAGTCTGCTGAGTGAATTAATACTATCAGCTTCA 988
DB 790 AAGTGACGAGCAGCTGCAACCGGAGTCTCTTGAAGGGGAAGAGTTGAACATTTGCTTCA 849
QY 989 ATCCAACTTACCTTATGAGCTTTTAAAGCTATTTAAAGTGAAACAGATAAAATTC 1046
DB 850 GCGGAAATATATGATGAGCGCTTGGCGGCGCTTGATGGAACAGACATTTCAATCA 907

RESULT 35
US-10-673-127-173
; Sequence 173, Application US/10673127
; Publication No. US20040081995A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 2221/1030
; CURRENT APPLICATION NUMBER: US/10/673,127
; PRIOR FILING DATE: 2003-09-26
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 173
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-10-673-127-173

Query Match 7.8%; Score 88; DB 16; Length 992;
Best Local Similarity 46.5%; Pred. No. 2.4e-08;
Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;

QY 214 CCAGAGCTATTTATGAGCTAGTTTTTTTATTAATATTTTCAAGTTTGCAGAT 273
DB 73 CCGGAGGAGCATGCTACTGCGCGCGCTTTTCTCTGAAATCGTGAAGGAACTCGCGCA 132
QY 274 ATTAGTATAATGTTAAAGAAATTAACAACACCAAGTTGTTTAAACAGTGTAAATCA 333
DB 133 CAACCGGTGGAAATCGAACCGGAGACAACTTTTGTGAGCATCATCCGCTCGGGGCACTCA 192
QY 334 GAGATTACTTTAAAGGAAAAGATGTTGACCAAGTATCTCTGCTTACAAAGAGTATCAACA 393
DB 193 GAATTCGCGCTCAATGGGCTAAACCGCGAGATATCGCGCTCGCGCAATTAAGAA 252
QY 394 GAAATCTCTGATTTTAAACAAAATTTATGAGTCTATTTATGCTGGAACAGCTTTT 453
DB 253 GAAACGCTGTTTCAAAATCCGCGCTGATTTATTTGAAACCGTGTATTCGGCAACGCTGTC 312
QY 454 GCAGCCAGTTTACAAGAAAGTCTCTATTTTAAACAGGATTCATTTGTTATTAAGTAA 513

DB 313 GCCGTTTCTACATCGGAAACCGCGCCCAATCTTGACAGGTGTCAACTGGAAAGT---TGAA 369
QY 514 CATAAAGATTTAAAGCAGTAGCGACTGACTCTCATCGTATGAGCCAAACGTTTAACTACT 573
DB 370 CATCGGAGCTTGTCTGCACAGCGACGACAGTCTATCGCTTAGCCATCGGCAAGTGAAA 429
QY 574 TTGACA---ATACTTCAGCAGATTTGATGTTAGTCTTCTTCAAGTAATCTTTGAGAAA 630
DB 430 ATTGAGTCGGAAAATGAAGTATCATCAACGCTCGTCATCCCTGGAAAAAGTCTTTAATGAG 489
QY 631 TTTTCAGCAGTATTTACAGATGATATTGAGACCGTTGAGGTATTTTCTCACCAGCCAA 690
DB 490 CTCAGCAAAATTTTGGATGACGCCAACACCCGGTGGACATCGTCATCAGCCCAATCAA 549
QY 691 ATCTGTTTCAAGTAGAACACATTTCTTTTATACAGCTCTTTAGAGGAAATTAATCC 750
DB 550 GTGCTATTAAAGCGGAGCACCTTCTCTCTTTTCCCGCTGCTTGACGGCAACTATCCG 609
QY 751 GATACAGACGCTTATTAAATGACAGAAATTTGAGACGGAGTTGTTTCAATACCAATCC 810
DB 610 GAGACGCGCCCTTGAATTTCCAAACAGAAAGCAACCATGATCGTCAATGCAAAAGAG 669
QY 811 CTTCGCCAGCTATGGAACGTCCTTCTTGATT--TCTAATGCTACTCAAAATGGTACTG 868
DB 670 TTCTGCGAGCAATCGACCGAGCGTCTTGCTGCTCGAGAAGGAAGCAACGTTGTG 729
QY 869 TTAAGCTTGAGATTACTCAAAATCATATTTTCAGCTCATGTTAACTCACTGAGTTGGTA 928
DB 730 AAATGACGAGCTTCTGAGGAATGCTCGAAATTTCTCGATTTCTCCGAGATCGGGA 789
QY 929 AGGTAACGAGGATTTAGATATTTAGTCTAGTCTGCTGAGTGAATTAATACTATCAGCTTCA 988
DB 790 AAGTGACGAGCAGCTGCAACCGGAGTCTCTTGAAGGGGAAGAGTTGAACATTTCTGTTCA 849
QY 989 ATCCAACTTACCTTATGAGCTTTTAAAGCTATTTAAAGTGAAACAGTAAAAATTC 1046
DB 850 GCGGAAATATATGATGAGCGCTTGGCGGCGCTTGATGGAACAGACATTTCAATCA 907

RESULT 36
US-10-670-817-173
; Sequence 173, Application US/10670817
; Publication No. US20040106137A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 2221/1030
; CURRENT APPLICATION NUMBER: US/10/670,817
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: US/09/716,964
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 173
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-10-670-817-173

Query Match 7.8%; Score 88; DB 17; Length 992;

Db 790 AAGTACGGAGCAGCTGCAACCGAGTCTTTGAAGGGGAAGAGTTGAACATTTTCGTTC 849
Qy 989 ATCCAACTTACCTTATTTAGAGTCTTTAAAGAGTATTAAAGTGAACAGAGTAAATTC 1046
Db 850 GCGCGAATATATGATGAGCGGTTGCGGGCGCTTGATGGAACAGACATTTCAATCA 907

RESULT 38

US-10-671-207-173
; Sequence 173, Application US/10671207
; Publication No. US20040197796A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmski, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/671,207
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US/09/716,964
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 173
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-10-671-207-173

Query Match 7.8%; Score 88; DB 18; Length 992;

Best Local Similarity 46.5%; Pred. No. 2.4e-08;
Matches 390; Conservative 0; Mismatches 440; Indels 8; Gaps 3;

Qy 214 CCAGGAGCTATTTTATAGAGCTAGTTTTTTTATTAATATTTTCAAGTTTGCAGAT 273
Db 73 CCGGGAGCATCGTACTGCAGCGCGCTTTTCTGAAATCGTGAATACTGCGCA 132
Qy 274 ATTAGTAAATGTTAAAGAAATGGAACCAACCAAGTTGTTTAAACAGTGGTAAATCA 333
Db 133 CAAACGGTGGAAATCGAAACGGAAGCAACTTTTTTGACGATCATCGCTCGGGGCACTCA 192
Qy 334 GAGATTACCTTAAAGGAAGATCTTGACAGTATCTCTGCTCAAGAGTATCAACA 393
Db 193 GAATTCGGCTCAATGGGGTAAACCGGACGATATCGCGCTCGCGCAATTTGAAGAA 252
Qy 394 GAAATCTCTTGAATTTTAAACAAATTTTGAAGTCTATTATTGCTGAAACAGCTTTT 453
Db 253 GAAACGTTGTTTCAATCCGCTGATTTTGAACCGTATTCGGCAACGTTGTTTC 312
Qy 454 CGAGCGAGTTTCAAGAAAGTCTGCTCTATTTTAAACAGGAGTTTCATTTGTTATTAAGTAA 513
Db 313 GCGGTTTCTACATCGGAAACGCGCCCAATCTTGACAGGTGTCAACTGGAAAGT---TGAA 369
Qy 514 CATAAGATTTTAAAGCAGTACGCTGCTCTCATCTGATGAGCAACGTTTATCACT 573
Db 370 CATGCGAGCTGTCTGACACGCGCCGACGATCTGCTTAGCCATGCGCAAGTGA 429
Qy 574 TTGGACA---ATACTTCAGCAGATTTGATGGTAGTTCTTCCAAAGTAAATCTTTTGAGAGAA 630
Db 430 ATTGAGTCGGAATGAAGTATCATACACGTCGTCATCCCTCGGAAGAGTCTTTAATGAG 489

Qy 631 TTTTCAGCAGTATTTTACAGATGATATTGAGACCGTTGAGGTATTTTCTCACCAGCCAA 690
Db 490 CTCAGCAAAATTTTGGATGACGGCAACCAACCGGTGGACATCGTCATGACAGCCCAATCAA 549
Qy 691 ATCTTTTCAAGTGAACACATTTCTTTTATACACGCTCTTGAAGGAAATTTATCCC 750
Db 550 GTGCTATTTAAGGCGGAGCACCTTCTCTTTTCCCGGCTGCTTGACGGCACTATTCG 609
Qy 751 GATACAGACCGTTTATTAATGACAGAAATTTGAGACGGAGGTTGTTTTCAATAACCAATCC 810
Db 610 GAGACGGCCGCTTGATTTCCAAACAGAAAGCAACGACCATGATCGTCAATGCAAAAGAG 669
Qy 811 CTTCGCCAGCTATGGAACGTCCTTTCTTGATT--TCTAATGCTACTCAAAATGCTACTG 868
Db 670 TTTTCGAGGCAATCGACCGGCTCTTGTGTCGAGAGGAAAGCAACGCTTGTG 729
Qy 869 TTAAGCTTGAGATTAATCAATATTTTCAAGCTCATGTTAACTCACCTGAGGTTGGTA 928
Db 730 AAATGACGACGCTTCTGAGGAGATGCTCGAAATTTCTTCGATTTCTCCGAGATCGGA 789
Qy 929 AGGTAACGAGGATTTAGATATTTGTTAGTCTAGTCTGTTAGTGAATTAATACTATCAGCTTCA 988
Db 790 AAGTACGAGCAGCTGCAACGAGTCTCTTGAAGGGGAAGAGTTGAACATTTTCGTTCA 849
Qy 989 ATCCAACTTACCTTATTTAGTCTTTTAAAGCTATTTAAAGTGAACAGTAAAAATCA 1046
Db 850 GCGGAAATATATGATGAGCGCTTTCGCGGCGCTTGTGGAACAGACATTTTCAAAATCA 907

RESULT 39

US-10-282-122A-15836
; Sequence 15836, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 15836
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-15836

Query Match 6.5%; Score 74; DB 16; Length 1101;
Best Local Similarity 46.0%; Pred. No. 2.1e-05;
Matches 503; Conservative 0; Mismatches 555; Indels 36; Gaps 6;

```
QY 4 ATTCATTTTCAATTAATCGCACATATTTATTCATCAATAAAAAATTTGAAGTCACTTCTACA 123
Db |||||
QY 1 ATCAATTTTATATGTGAAAAAATATATTACAAGAGCAATTTATCAGCTCAAAAAGCT 60
Db |||||
QY 64 ATTAGACATAAAATGCCATTCCTATCTTTCATCAATAAAAAATTTGAAGTCACTTCTACA 123
Db |||||
QY 61 GTTACAGGAAATCAACAATGCCAGTATTACAAGGTATCTCATGAGTGTTCAAAATAAC 120
Db |||||
QY 124 GGAGTAACCTTTAACAGGCTTAAACGGTCAAAATATCAATTTGAAACACACTATTCCTGTAAGT 183
Db |||||
QY 121 GAAATTAATTTTAAATTTGGTCTGATATAGATTAAAGTATTGAACTAAATAAATGAGAA 180
Db |||||
QY 184 AATGAAATGCTGGTTTCTAATTAATCTCTCCAGGAGCTATTTATTAGAGCTAGTTT 243
Db |||||
QY 181 GTTTTAGAAGAAGTAAAGTTGTT-----TTAGATGCTAGGCTT 219
Db |||||
QY 244 TTTATTATATATTTCAAGTTTCCAGATATTAGTATAAATGTTAAAGAAATGAAACA 303
Db |||||
QY 220 TTAAGTGAATAATATAGAAAGCTGCCATAATCCAAAGTTGAGATACAACTATTGAGAT 279
Db |||||
QY 304 CACCAAGTGTGTTTAAACAGTGGTAAATCAGAGATTACCTTAAAGGAAAGATGTTGAC 363
Db |||||
QY 280 AATTGTGTGAAATAACTTGTATTAATCAAGCTAACACTTGTGTATCTAAATCCAAT 339
Db |||||
QY 364 CAGTATCTCTGCTACAGAGTATCAACAGAAATCTCTTGTATTTTAAACAAATA 423
Db |||||
QY 340 GATTTTCCAGCTTACAGAGATAGAGAAACTCAATCTTCAAAATAATCAAAAACA 399
Db |||||
QY 424 TTGAAGTCTATTTTGTGTAACAGCTTTTGGAGCCAGTTTACAGAAAGTGTCTTAT 483
Db |||||
QY 400 TTAATAACTATGATAAAGGTACTATTTTGTCTTATTTGCTCAGATGAAACAGGCTATA 459
Db |||||
QY 484 TTAACAGGAGTTCATATTTGTTAAGTAAATCAATAAGATTTTAAAGCAGTACGAGCTAC 543
Db |||||
QY 460 TTAACAGGCGGTTTATTTGAAATA---AAAGATAGCAAAATTAATTTAGTTGCTATTGAT 516
Db |||||
QY 544 TCTCATCGTATGAGCAACGTTTAACTCTTGGACATATCTTACAGAGATTGATGTA 603
Db |||||
QY 517 GGGTATAGCTAGCTTTAG---AAGCCATACATAGATAACGAACCCAGTATAAATGCA 573
Db |||||
QY 604 GTTCTTCCAAAGTAAATCTTTGAGAGAAATTTTTCAGCAGTATTTACAGATGATTTGAGACC 663
Db |||||
QY 574 GTAATTTCCAGGAAAAACATAAATGAAGTTATTAATAATTTTAGAGGATGTTGTA---T 630
Db |||||
QY 664 GTTGAGGTATTTTCTCCAGAGCAATCTTGTTCAGAGTGAACACATTTCTTTTAT 723
Db |||||
QY 631 GTAAATATATACATTTACTTACCATATATTTCAATCTTGGAAATCTAAGATAATA 690
Db |||||
QY 724 ACAGCCTCTTAGAAGGAAATTTATCCGATACAGACCGTTTATTAAATGACAGATTTGAG 783
Db |||||
QY 691 TCAAGACTACTTGAAGGTGAATTTATAAAATATAATTAATCAATATACCTGAGATATCAAC 750
Db |||||
QY 784 ACGGAGGTGTTTTCATACCCATCCCTTCGCCACGCTATGGAAGTGTCTTTGAT 843
Db |||||
QY 751 TTAATATTTGGCTAGAAAAGAGAAATTTGTTAGACTGTCATAGAGAGCTTCACTTATG 810
Db |||||
QY 844 TCTAATGCTACTCAAAATGTTACTGTTAAGCTTTCAGATTTACTCAAAATCATATTTTCACT 903
Db |||||
QY 811 GCTAAGGATGAAATAACAATTTGATAAATTAGACAT---CGAGATGATGTAATGATT 867
Db |||||
QY 904 CATGTTAACTCAGCTGAGTTGGTAAAGTAAACAGGATTTAGATTTGTTAGTCACTCT 963
Db |||||
QY 868 ATAAGTCCAAATCTCAATTTGGGAAACGTAAGAGAAATAAATAATTTTGCA---A 924
Db |||||
```

```
QY 964 GGTAGTGATTTAACCTATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTAAAAGCTATT 1023
Db |||||
QY 925 GGTGAGCCTTTGAAAATTTGCATTTAACTCAAAATATTTAATAGATGTCTTTAAAATCATG 984
Db |||||
QY 1024 AAAAGTGAACAGTAAAAATTCATTTCTTATCACCAGTTCGACCATTCACCTTAACACCA 1083
Db |||||
QY 985 AATCAGGAAGAATTTGTAATGNACTTTTCAAGTAGCATAGTCTTGCATTTATTTAAAAAT 1044
Db |||||
QY 1084 GCGGATGAGGAAGA 1097
Db |||||
QY 1045 AAGGAAATGATGA 1058
Db |||||
```

RESULT 40

US-09-974-300-6041
; Sequence 6041, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6041
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6041

Query Match 6.0%; Score 68.6; DB 9; Length 486;
Best Local Similarity 48.8%; Pred. No. 0.00019;
Matches 216; Conservative 0; Mismatches 224; Indels 3; Gaps 1;

```
QY 664 GTTGAGGTATTTTCTCACCAGCCAAATCTTGTTCAGAGTGAACACATTTCTTTTAT 723
Db |||||
QY 15 GCTGAAATTTGTTAAACCGAGATCAATCTGTTCAAACCTCGGAACTTGTGTTTTC 74
Db |||||
QY 724 ACAGCCTCTTGAAGGAAATTTCCCGATACAGACCGTTTATTAAATGACAGAAATTTGAG 783
Db |||||
QY 75 TCAAGACTTCTTGAAGGGAAGTATCCAGTAACGTCGAGCATGATACCGACTCAGTTCAA 134
Db |||||
QY 784 ACGGAGGTGTTTTCATACCCAAATCCCTTCGCCACGCTATGGAACGCTTCTTGTATT 843
Db |||||
QY 135 ACGTCTTGTGCTAATAACCAAGAAATTTGCTCATAGCTTGAACCGGCAATGTTGTTG 194
Db |||||
QY 844 TCTAATGCTACTCAAAATGTTACTGTTAAGCTTGAGA---TTACTCAAAATCATATTTCA 900
Db |||||
QY 195 TCAAGGGAAGCAAAAACAATGTCAATTAACGTCAAAACACTTGGCGACACCAATTTGAA 254
Db |||||
QY 901 GCTCATGTTTAACCTCACCTGAGTTGTTAAGTAAACGAGGATTTAGATATTTGTTAGTCA 960
Db |||||
QY 255 ATTAGTCCACATCAAGAATAAGGAAAGTAGCGGAAAAATGGCCACTGACCAATTT 314
Db |||||
QY 961 TCTGTTAGTATTAATATCAGCTTCAATCCAACTTACCTTATTGAGTCTTTTAAAGCT 1020
Db |||||
QY 315 GAAGGGATGAATTCGGATTTGTTTAAATGGGAAAAATATTTCTGATGCTTGAAGTG 374
Db |||||
QY 1021 ATTAAGTGAACAGTAAAAATTCATTTCTTATCACCAGTTCGACCATTCACCTTAACA 1080
Db |||||
QY 375 GTTGACAGCATGAGTCCAAATCTTTTGAAGGTCGATGAGTCCCTCGTTATTTCG 434
Db |||||
QY 1081 CAGGCGATGAGGAAGAAAGTTT 1103
Db |||||
QY 435 CCAATTGATTACAGACCACTATT 457
Db |||||
```

Fri Feb 4 10:26:48 2005

us-10-048-071-27.rnpb

Page 33

Search completed: January 31, 2005, 13:50:53
Job time : 508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2005, 16:54:57 ; Search time 40 Seconds
(without alignments)
626.705 Million cell updates/sec

Title: US-10-048-071-28
Perfect score: 1863
Sequence: 1 MIQFSINRTLFHALNTTKR.....LTPGDEESFIQLTPVTN 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCITUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1389	74.6	378	US-09-583-110-4405	Sequence 4405, Ap
2	863.5	46.3	384	US-09-107-532A-4556	Sequence 4556, Ap
3	748.5	40.2	385	US-09-134-001C-5166	Sequence 5166, Ap
4	431	23.1	181	US-09-134-000C-5039	Sequence 5039, Ap
5	350.5	18.8	368	US-09-252-991A-18904	Sequence 18904, A
6	315	16.9	383	US-09-328-352-5969	Sequence 5969, Ap
7	313	16.8	374	US-09-489-039A-10579	Sequence 10579, A
8	307	16.5	375	US-09-543-681A-7400	Sequence 7400, Ap
9	249.5	13.4	375	US-09-818-780-23	Sequence 23, Appl
10	232	12.5	303	US-09-198-452A-357	Sequence 357, App
11	131.5	7.1	141	US-09-134-000C-5040	Sequence 5040, Ap
12	117.5	6.3	977	US-09-248-796A-15579	Sequence 15579, A
13	115.5	6.2	800	US-08-776-265-3	Sequence 3, Appli
14	115.5	6.2	800	US-09-398-184-3	Sequence 3, Appli
15	114.5	6.1	823	US-09-248-796A-16699	Sequence 16699, A
16	113.5	6.0	922	US-09-883-134-9	Sequence 9, Appli
17	112.5	6.0	470	US-09-248-796A-23131	Sequence 23131, A
18	112.5	6.0	619	US-09-066-046-2	Sequence 2, Appli
19	109.5	5.9	1753	US-09-248-796A-19154	Sequence 19154, A
20	109	5.9	532	US-09-710-279-546	Sequence 546, App
21	107.5	5.8	569	US-09-248-796A-16697	Sequence 16697, A
22	107	5.7	563	US-09-134-001C-3172	Sequence 3172, Ap
23	106.5	5.7	1155	US-09-543-681A-6286	Sequence 6286, Ap
24	104.5	5.6	1726	US-09-700-227-2	Sequence 2, Appli
25	103.5	5.6	564	US-09-107-532A-5248	Sequence 5248, Ap
26	103.5	5.6	920	US-09-463-402-6	Sequence 6, Appli
27	103.5	5.6	921	US-09-889-572-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-583-110-4405
; Sequence 4405, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4405
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4405

Query Match	74.6%	Score 1389	DB 4	Length 378
Best Local Similarity	72.2%	Pred. No. 6.7e-125	Mismatches 55	Indels 0
Matches 273	Conservative 50			Gaps 0
QY	1	MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGTTLTSGNGQISIENTIPV	60	
Db	1	MHFSINKNLFQALNTTKRAISSKNAIPILSTKIDVTNEGITLIGSNGQISIENTIFISQ	60	
QY	61	SNENAGLITSPGAILLEASFPINISLDPISINKEIQHVLTSGSKSITLKGKDV	120	
Db	61	KNEADGLITSLGSLLEASFPINVSULPDVTLTKEIQNQIVLTSGSKSITLKGKDS	120	
QY	121	DOYPLEQVSTENPLILTKLLKSIATFAASTQESRPILTGVHIVLSNKHDKAVAT	180	
Db	121	EOYPRIQEASATPILTKLLKSIINETAFAASTQESRPILTGVHIVLSNKHDKAVAT	180	
QY	181	DSHRMSQRLITDNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF	240	
Db	181	DSHRLSQKLTLEKNSDDFVVPISRLREFSAVFTDDIETVEIFFANQILFRSEHISF	240	
QY	241	YTRLEGNYPTDRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLEITQNHIS	300	
Db	241	YTRLEGNYPTDRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLEIKDGVVS	300	
QY	301	AHVNSPEVKVNEEDLDIVSQSGSDLTISFNPTLIESLKAIKSETVKIHFSPVPTLT	360	
Db	301	AHVNSPEVKVNEEDLDIVSQSGSDLTISFNPTLIESLKAIKSETVKIHFSPVPTLT	360	

Db 301 AHVHSPVGVKNEEDTDQVTGDLTISFNPTYLIDSLKALNSEKVTISFISAVRPFTLV 360
QY 361 PGDEEESFIQLITPVRTN 378
Db 361 PADTDEDFMQLITPVRTN 378
RESULT 2
US-09-107-532A-4556
; Sequence 4556, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...384
; SEQUENCE DESCRIPTION: SEQ ID NO: 4556:
US-09-107-532A-4556
Query Match 46.3%; Score 863.5; DB 4; Length 384;
Best Local Similarity 46.6%; Pred. No. 2.1e-74;
Matches 176; Conservative 72; Mismatches 129; Indels 1; Gaps 1;
QY 1 MIQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNQISIENTIPV 60
Db 8 INKVLNRAFQWELQTVQRAISSKTTIPILGVKLTITQEGTLTGSNADISIEFLSV 67
QY 61 SNEAGLLITSPGAILLEASFINISSLDPDISINVKIEIQHVLTSGSKSEITLKGKD 120
Db 68 ENEKANMQIESTGSIVLQARFFSEIIRLPETFTILEVLENQVAITSGKANFVINGLDA 127
QY 121 DQYPRLOEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNKHDKFVAT 180
Db 128 DNYPHLPVNVESHNQMKLPVHVLTKLINETVFAVSQHSRPILTGVHIFILSDN-SLLAVAT 186

QY 181 DSHRMSQRLITLNDNTSADLMVVLPSKSLRFSVAFTDDIETVEVFFSPQILFERSEHSF 240
Db 187 DSHRLSORVIPVEQAAADHFDIVIPGKSLIELSLTNEEELVISEMENQVLFTETMYF 246
QY 241 YTRLLEGNYPDTDRLLMTETETEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
Db 247 YSRLEGNYPDTNRLIPSSFNTEVEFVSPFLAAIERASLLSHEGRNNIVRLSRPDAVV 306
QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSLPVRPFTLT 360
Db 307 LYGNSPRIGKVEESLYTASSGDLDFISFNDYMKALRAPGDMWSIKVKFISAIRPFTLE 366
QY 361 PGDEEESFIQLITPVRTN 378
Db 367 PTEGQVFIQLITPVRTN 384
RESULT 3
US-09-134-001C-5166
; Sequence 5166, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5166
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5166

Query Match 40.2%; Score 748.5; DB 3; Length 385;

Best Local Similarity 40.5%; Pred. No. 2.4e-63;
Matches 153; Conservative 86; Mismatches 136; Indels 3; Gaps 3;

QY 1 MIQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNQISIENTIPV 60
Db 9 MMEFTIKRDYFINQLNDTLKAI SPRTTLPILTIKIDAKENEVILTSGSDSEISIEITPK 68
QY 61 SNEAGLL-ITSPGAILLEASFINISSLDPDISINVKIEIQHVLTSGSKSEITLKGKD 119
Db 69 QVDGBEIVEITETGTVLPGRFVVDIIKKLPKKEVKLSTNEQFQTLITSGHSEFNLSGLD 128
QY 120 VDOYPRLOEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNKHDKFVAT 179
Db 129 PDQYPLLEVRSDDAIQLSVKLVNIIAQTFAVSTETRPVLGTGVNWLIOQN-ELICTA 187
QY 180 TDSHRMSQRLITLNDNTSADLMVVLPSKSLRFSVAFTDDIETVEVFFSPQILFERSEHS 239
Db 188 TDSHRLAVRKLQLEDESENKVNIIIPGKALSINKIMSDSDDEDIDIFFASNQVLFRVGNIN 247
QY 240 FYTRLLEGNYPDTDRLLMTETETEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 299
Db 248 FISRLLEGHYPTDTRLFPENYEIKLGINNGDFYHAIDRASLLAREGGNNVRLKLTGNELV 307
QY 300 SAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSLPVRPFTLT 359
Db 308 ELSSTSPRIGTVKEVNANDVEGGNKLISFNKYMMDALKDAIDNDEVEFEFGTKPFTL 367
QY 360 TPGDEEESFIQLITPVRT 377
Db 368 KPCKD-DDSVTQLILPRT 384

RESULT 4

US-09-134-000C-5039
 ; Sequence 5039, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 5039
 ; LENGTH: 181
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-5039

Query Match 23.1%; Score 431; DB 4; Length 181;
 Best Local Similarity 47.5%; Pred. No. 2.2e-33;
 Matches 86; Conservative 35; Mismatches 60; Indels 0; Gaps 0;
 QY 198 DLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISFYTRLLEGNYPTDRLIM 257
 DB 1 DFNIVPGKSLTELSRLTNESEVMSIMENQVLFKTEWTFYFGRLLGNYPTDRLIP 60
 QY 258 TEFTEVVENTOSLAHAMEAELISNATONGVVKLEITONHISAHVNSPEVCKVNEIDL 317
 DB 61 TSHNTQIEFYVELLSAERALLSHEGNNVRLISGSDSVVLYGNSPEIGKVEEALNY 120
 QY 318 VQSQSDLTISFNPTYLIESLKAISKSETVKIHFSLSPVFPFTLTPGDEBESFQLITPVRT 377
 DB 121 ENVSGEALDISPNYMDKALRAFGDMNITVKVFLSPIRPFLEPTETELDFQLITPVRT 180
 QY 378 N 378
 DB 181 N 181

RESULT 5

US-09-252-991A-18904
 ; Sequence 18904, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1997-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 18904
 ; LENGTH: 368
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-18904

Query Match 18.8%; Score 350.5; DB 4; Length 368;
 Best Local Similarity 24.4%; Pred. No. 3.8e-25;
 Matches 92; Conservative 95; Mismatches 177; Indels 13; Gaps 8;
 QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSGLTSGNSQISIENTIPVS 61
 DB 2 MHFTIQREALLKPLQVAGVVRRTQLPVLNVLVVEQQSLGTLTDELEVLRVILE 61
 QY 62 NENAGLLITSPGAILLEASFNNIISLP-DISINVKIEHQVLTSGKSEITLKGKDV 120

DB 62 D-----AAEPGEITVPARKLMDICKSLPNVDLDIR-VEOKLVKAGRSRFTLTSLPA 114
 QY 121 DQYPRLOEVSTENPLILTKLKSIIAETAFAASLQESRPILTVGHIVLSNHNKDFKAVAT 180
 DB 115 NDFPTVEGPGSLNFSIAQSKLRLIDRTSPAMAQDVRYYLNGM-LLEVNGGTLRSVAT 173
 QY 181 DSHRMSQRLITLNDNTSADL-MVVLPSKSLRSPSAVFTDDIETVEVFPSPQILFRSEHIS 239
 DB 174 DGHRLAMCSLDAQIPSDRHOVIVPRKGILELARLLTEQDGEVGIIVLGQHHIRATTGEPT 233
 QY 240 FVTRLLEGNYPTDRLAMTEFETEVVNTQSLRHAMERAFILSNATONGTVKLEITQNH 299
 DB 234 FTSKLVGDKFPDYERVLPRGGKLVGDRQQRREAFSRTAILSNKRYG-IRLQUSNGLT 292
 QY 300 SAHVNSPEVCKVNEIDLIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSLSPVFPFTL 359
 DB 293 KIQANNPEQEAEEVQ-VEYNGGNLEIGFNVSYLLDVLGVGTGTEQVRF-ILSDSNSSAL 350
 QY 360 TPGDEBESFQLITPV 376
 DB 351 VHEADNDDSAVVMEMR 367
 RESULT 6
 US-09-328-352-5969
 ; Sequence 5969, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 5969
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-5969

Query Match 16.9%; Score 315; DB 4; Length 383;
 Best Local Similarity 24.3%; Pred. No. 1e-21;
 Matches 95; Conservative 100; Mismatches 172; Indels 24; Gaps 12;
 QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSGLTSGNSQISIENTIPV 60
 DB 1 IVRLKIAKESLNLVSHVGVAVERRHTNLILSNVKIQNAQALITIGSDLELVASTAL 60
 QY 61 S-EGACL--EAGETTVPARKLMEICKSLPTAALIDLQITEDQRCILKSGNSRFLVLTLP 116
 DB 61 S-EGACL--EAGETTVPARKLMEICKSLPTAALIDLQITEDQRCILKSGNSRFLVLTLP 116
 QY 120 VDQYPRLOEVSTENPLILKT-KLKSIIAETAFAASLQESRPILTVGHIVLSNHNKDFKAV 178
 DB 117 AEDYPLLTENSQGTQVQVQRELKRLFEKTAFAVQDVRFYLTGTLLEIDENQ-LRAV 175
 QY 179 ATDSHRMSQRLITLNDNTSADLM-VVLPSKSLRSPSAVFTDDIETVEVFP-----S 227
 DB 176 TTDGHRALCELASSTSSQLVQAIIVPRKAVGELQRLLSIEDQTLTLLIGRELLNVTINT 235
 QY 228 PSQILFRSE-HISFYTRLLEGNYPTDRLAMTEFETEVVNTQSLRHAMERAFILSNATQ 286
 DB 236 PSROKEQGDITVRETTKLIDGKFPDYRRVIRPGGDKHVLIGHDVFKQSLQRAVILSNEKL 295
 QY 287 NGTVKLEITONHISAHVNSPEVCKVNEIDLIVSQSGDLTISFNPTYLIESLKAISKSETV 346
 DB 296 RG-VFLFNQDLSQLRANNPQEADIEDLAIOYQS-APLEMSFNAQYLLDVLGVLDGDDV 353
 QY 347 KIHFLSPVFPFTL-TPGDEBESFQLITPV 376
 DB 354 NMSMTEANQSVLVQDPAHPDQTYV--VMPMR 382

```

US-09-543-681A-7400

Query Match          16.5%; Score 307; DB 4; Length 375;
Best Local Similarity 24.2%; Pred. No. 5.9e-21;
Matches      85; Conservative      85; Mismatches 165; Indels    16; Gaps     7;

QY   2 IQPSINRTLFIHALNTTKRAISTKNAIPILSSITKIEVTSVTGVLTCGNGOISIENTIPVS 61
DB   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~::|:|
       9 MKFIIEREQLLKPLQOVSGPLGRPTLPILGNLLLVKTENTLSLTGTDLMEMMARVSL 68
QY   62 NENAGLLITSPGAILLEASFFINISSLP----DISINVKEIEHQHVVLTSCKSEITLKGK 118
DB   69 QSH-----EIGATTVPARKFFDIWRGLPEGAISV---ELDGDRLLVRSGRSFSSTL 119
QY   119 DVQYPRLOEVSTENPELLILTKLKLSIAETAFASLQBSRPILTGVHIVLSNHKDPKAV 178
DB   120 PASDFPNLDWDQOSEVEFTLPQATLKRLIESTQSMAHQDVRYTLNGW-LPETENTEIRTV 178
QY   179 ATDSHRMSQRLLTLDNITSADLMVLVPSKSLRFPESAFTDDIET-VEVFPSPOILPRSEH 233
DB   179 ATDGHRLAVCAMDIGOSLPCHSVIVPRKGVIELMRLLDGSGESLLQLQTGSNNLRHAVGD 238
QY   238 ISFYTRLLEGNYPDTRDLLMTTEFETEFVFNFTQSLRHAMERAFLLISNATONGTVKLEITON 297
DB   239 FIETSKLVDPGRPPDYRRVRLPKNPCTKVIAACDILKQAFSAAILSNKEFRPG-VRIINTNG 297
QY   298 HISAHVNSREVGKNEDLIIVSOGSDLTISFNPTVYLIESLKAIKSETVKI 348
DB   298 QLKITANNPEQEAEISVDVQYO-GBEMEIGFNVSLLDVDLNTLKCEVKL 347

RESULT 9
US-09-818-780-23
; Sequence 23, Application US/09818780
; Patent No. 6677146
; GENERAL INFORMATION:
; APPLICANT: McHenry, Charles
; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
; FILE REFERENCE: 1794.0030004
; CURRENT APPLICATION NUMBER: US/09/818,780
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/192,736
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Thermus thermophilus
US-09-818-780-23

Query Match          13.4%; Score 249.5; DB 4; Length 375;
Best Local Similarity 23.1%; Pred. No. 2e-15;
Matches      90; Conservative      85; Mismatches 185; Indels    29; Gaps    14;

QY   2 IQPSINRTLFIHALNTTKRAISTKNAIPILSSITKIEVTSVTGVLTCGNGOISIENTIPVS 61
DB   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~::|:|
       1 MNITVPKLLSDQLSLLERIVPERSANPLYTLGLYAEBCALLFGTNGEVDLEVRPAE 60
QY   62 NENAGLLITSPGAILLEASFFINISSLPD----ISINVKEIEHOHVVLTSCKSEITLKG 117
DB   61 AQ-----SLPVLVPAQPFFOLVNSLPDDLVALGLASEPGGGQLSELSSGRFRTL 113
QY   118 KDVDQYPRL--QEVSTENPLLILTKUL-----LKSIIAETAFASLQBSRPILTGVIHLVSN 171
DB   114 APAEGYPELLVPEGEDKGAFFLTRMPSGELVKALTHRVYAASNEEYRAIFRGVOLFEFS- 172
QY   172 HKDPKAVATDSHRMSQRLLTLDNITSADLMVLVPSKSLRFPESAFTD-DIETVEVFPSPQ 230
DB   173 PQGFRAVASGYRLAYLDVLPLO-GFOAKAVVPARSVDENVMLVKAGDAEADLALGEV 231
QY   231 ILPRSE---HISFYTRLLEGNYPDTRDLLMTTEFETEFVFNFTQSLRHAMERAFLLISNATON 287
DB   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~::|:|

```

Db 232 LALALEGGGVRMALRLMEGEPDQVRVPIQBFALRVQVEGALREAVRVSLSD-RON 290
QY 288 GTVKLEITQNHISAHVNSPEVGVKVNEDLDIVSQSGDLTISFNPTLYLIESLKAISKSETVK 347
Db 291 HRVDLLLEGRILLSAEG-DYCKGQEVF-AQVEGPGMAVYNAVRYLLEALAPV-CDRAH 347
QY 348 IHFLSPVRFTTTPGDEESFQLITPVR 376
Db 348 LGISGTSP-SLIWGD-GEGRVAVVPLR 374

RESULT 10
US-09-198-452A-357
; Sequence 357, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 357
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-357

Query Match 12.5%; Score 232; DB 4; Length 303;
Best Local Similarity 25.4%; Pred. No. 6.6e-14;
Matches 67; Conservative 63; Mismatches 124; Indels 10; Gaps 7;

QY 115 LAGKVDQVPRLOEVSTENPLILTKLKSIIAETAFASLOESRPILTGHVILSNKD 174
Db 44 LLSMEKEDFPMLPDQNALRFLPABQLMORTSFVRSRESRVLTGVLILAANGV- 102

QY 175 FKAVATDSHRMSQ--RLITDNTSADLMVLPKSLRFPESAVFTDDIETVEVFPSPQIL 232
Db 103 ATIVGDGRLAKIDAEVTLDSFG-EVLIIPKAVEELIKMSDEGEAT-IFLDQDKIA 160

QY 233 FSEHISFTRILEGNYPTDRLMTTEFEVFNQSLURHAMEAPLISNATQNGTVKL 292
Db 161 VECNTLLITLKLSEFPDFSPVISTESNVKLDHREELITLLKQVALFTNESSH-SVKF 219

QY 293 EITQNHISAHVNSPEVGVKVNEDLDIVSQSGDLTISFNPTLYLIESLKAISKSETVKIHL 352
Db 220 SFLPGELTUTANCKRVGKGVSM-AVNYSGELLEIAFPFFLDILKHSKDELVLGSD 278

QY 353 VPRPFTLPGDEESFIQLITPVR 376
Db 279 SYNPGIIT--DSASGLFVIMPWR 299

RESULT 11
US-09-134-000C-5040
; Sequence 5040, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5040
; LENGTH: 141
; TYPE: PRT

; ORGANISM: Enterococcus faecalis
US-09-134-000C-5040

Query Match 7.1%; Score 131.5; DB 4; Length 141;
Best Local Similarity 32.6%; Pred. No. 8.5e-05;
Matches 45; Conservative 16; Mismatches 46; Indels 31; Gaps 5;

QY 71 SPGAILLEASFFINISSLPDISINVKEIQHOV-----VLTSG---KSEITLKG 117
Db 11 SHGSIVVTSRFFWGIIRRLPEDMTMEVLNVLGGPFRGGKIPVGGGEGKREIPPPP 70

QY 118 KDVDQVPRLOEVSTENPLILTKLKSIIAE-----TAFASLOESRPILTGHVILSNH 172
Db 71 PFCGWAPPRE-----LPGSFGEKNIFTRGGGFMPSRPILTGVHFILENQ 118

QY 173 KDFKAVATDSHRMSORLI 190
Db 119 K-LLAVATDSHRLSORVI 135

RESULT 12
US-09-248-796A-15579
; Sequence 15579, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15579
; LENGTH: 977
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15579

Query Match 6.3%; Score 117.5; DB 4; Length 977;
Best Local Similarity 20.8%; Pred. No. 0.045;
Matches 93; Conservative 69; Mismatches 170; Indels 115; Gaps 16;

QY 5 SINRPLFIHALNTRKRAISTKNAIPILSIIKIEV-----TSGVTILTGSNGQI---SIE 55
Db 429 SIFPLTLTAIAAPTSTAMST-----DLSSSVNIPVETSSNESSITLASSESIAVQTSIE 484

QY 56 NTPVSN-----NAGLLITSPGAILLEASFFINISSLPDISINVKEIE 100
Db 485 --IPVSSHISANQESSIGESSVNGSTMISSSALESASEFTIAQSSDVESYSSSHSIT 542

QY 101 QHVVLTSG---KSEITLKGKDQVQPRLOEVSTENPLILTKLKLK-----SIIAETAPA 152
Db 543 HESSAITSSSTTIETHVTESVVDSEF-----TSDSLFVSETQDPKASSFSIVTPSDSI 596

QY 153 ASLOESRPILTGHVILSNHDKFKAIVATDSHRMSORLITDNTSADLMVLPKSL---R 209
Db 597 IVNESSIPELSTVGAVSTSTSSSISNSYPPSTHISKVSSSAEPTVPTPSESIIINI 656

QY 210 EPSAVFTDIE-----TVEVFPSPQILFRSEHI 238
Db 657 SEPAVSTDSVQSTPEQPMSSVIAESLISSTVSALELSTLMELSHESILLAKOSETI 716

QY 239 SFYTLLEGNYPD-----TDRLLMTTEFEVFNQSLURHAMEAPLIS-----282
Db 717 N-----QSESNYISETSAVPTAKSTISDGLTSTWESSSLVNLSTSVVVMSESEITPIV 772

QY 283 -----NATQNGTVKLEITQNHISAHVNSPEVGVKVNEDLDIVSQSGDLT---ISFNPTLYIE 336
Db 773 HTSESATESAESSDIQTSTTWTTKSDGFSVETSGVISQSGTSFTTATFPPLY---829

```

QY 337 SLKAISKETVKIHLSPVPRPFTLTGCD 363
Db 830 -----TSSDITTEFIS---TWATNSD 848

RESULT 13
US-08-776-265-3
; Sequence 3, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-776-265-3

Query Match 6.2%; Score 115.5; DB 3; Length 800;
Best Local Similarity 18.9%; Pred. No. 0.051;
Matches 89; Conservative 86; Mismatches 174; Indels 123; Gaps 20;

QY 2 IQFSINRTLFTHALNTTK-----RAISTKNAIPILSSIKIEVTSTGV---TLTG 47
Db 260 IPYEVNKGSLVKRIDELRAADKVDGIVEVREDTDTGLRIALIELKDVNSSEIKNYLYKN 319
QY 48 SNGQISIE-NTIPVSNENAGLITSPGAILLEASFF---INISLDPDISINVKIEIQHQ 103
Db 320 SDLQISVFNFNVAISDGRPKLM---GIRQIDSILNQHIEVAVNRTKFDLDAEKRMHI 375
QY 104 V-----VLTSGKSEITLKGKDQVDQVPLQEVSTENPLILTKLKLSIAE 148
Db 376 VEGLIKALSILDKVIELIRSSKNKDKAKENLIEVYEFEEQAEAVMLQLYRLNTLTL-- 433
QY 149 TAPRASLOESPIITGVHIVLSNH-----KDFKAVATDSHRMSQRLITLD 193
Db 434 VALEGHEKELBALIKQLRHILDNDALNVIKEELNEIKKKFK-----SERLSLIEABIE 488
QY 194 NTSADLMVVLPSK-----SLREPSAVFTDDI-----ET 221

Db 489 EIKIDKEVWVFSEEVILSMTRHGVIKRTISIRSFASGVEDIGLKDDSLKHQEVNTQDT 548
QY 222 VEVFSPSQIILFRSEHISFYTRLLEGNYPD-----TDRLLMTFETFEVVFVNTQSLRHAMER 277
Db 549 VLVFTNKGRLFPVH-----KLRIKMKELGQHVSVQIVPIE-EDEVVINVNEKDFNTD 602
QY 278 AFLISNATONGTVKLEITQNHISAHVNSPEVG-KVNEDLDIVSQ-----SGSD 324
Db 603 AFYVF-ATQNGMIKKSTVPLFKTTRPNKPLIATKVKENDDLISVMRFKQQLITVITNKG 661
QY 325 LTISFNPTYLIES-LKA--IKSETVKIHLSPVPRPFTLTGDEBESFIQLIT 373
Db 662 MSLYNTSELSDTGLRAAGVKSINLKVEDF-----VVMTEGVSENDTILMAT 708

RESULT 14
US-09-398-184-3
; Sequence 3, Application US/09398184
; Patent No. 6649394
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6649394el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,184
; FILING DATE: 17-Sep-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
; US-09-398-184-3

Query Match 6.2%; Score 115.5; DB 4; Length 800;
Best Local Similarity 18.9%; Pred. No. 0.051;
Matches 89; Conservative 86; Mismatches 174; Indels 123; Gaps 20;

QY 2 IQFSINRTLFTHALNTTK-----RAISTKNAIPILSSIKIEVTSTGV---TLTG 47
Db 260 IPYEVNKGSLVKRIDELRAADKVDGIVEVREDTDTGLRIALIELKDVNSSEIKNYLYKN 319

```


RESULT 17
 US-09-248-796A-23131
 ; Sequence 23131, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 23131
 ; LENGTH: 470
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-23131

Query Match 6.0%; Score 112.5; DB 4; Length 470;
 Best Local Similarity 21.3%; Pred. No. 0.041;
 Matches 76; Conservative 76; Mismatches 126; Indels 79; Gaps 20;
 QY 18 TKRAISTKNAIPILS-SIKIEVSTGTGTLTSGNGQISIENTIPVSNENAGLLI--TSPGA 74
 Db 1 TKRPGLSLSVINPTELESEIKPNPVL-----NQLSEH---PPNLTGVLVPTATAA 53
 QY 75 ILLEAS-----PFNIISLPDISINVKEIQHVLTSGKSEITLKGKVDQYPRL 126
 Db 54 TTTTSPTNHLHPYFVSLIPN-----PQHHPMLDTSQHQTSTTDNQDYSHL 103
 QY 127 QGVSTENPILK--TKLKSIIAETAF-----AASLQESRPLTGVHIVLSNKHDPKAV 178
 Db 104 -SMPOSTPLMKYDITTYVRLTETKYNYMLDLANELDIWKIIPSLCQISEN----- 156
 QY 179 ATDSHRMSQRLITLNTSADLMVLPKSLRFSFAVFTDDI--ETVEVFSPQLPRSE 236
 Db 157 --DSFLDCLMSCRSNVLN--DLTNEQLNKSQKNAVPISERIQOF-----E 203
 QY 237 HLSF-YTRLLEGNYPTDRLLMTETETEVFTQS--LRHAMER-----AFIIS 282
 Db 204 HLISIVLTLGLYLTTKVRLTDYH-KVIFNNQAKLSFHLVKIHTITSNKNNSAVLT 262
 QY 283 NATQNGT-VKLEITQNH-ISAHVNSPEVKYNEIDLIVSQSGSDI-----TISFNPTY 333
 Db 263 NAIQSITMLKPFIDKNYDFSEFRNIQKGRVTDLLEITYNSNLYSNPDISIYSTF 319

RESULT 18
 US-09-066-046-2
 ; Sequence 2, Application US/09066046A
 ; Patent No. 6204252
 ; GENERAL INFORMATION:
 ; APPLICANT: MURPHY, Cheryl
 ; STOREY, James
 ; BELTZ, Gerald A.
 ; COUGHLIN, Richard T.
 ; TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
 ; CORRESPONDENCE ADDRESSES: 65
 ; ADDRESS: HALE AND DORR LLP
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: United States
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/066,046A
 FILING DATE: 24-Apr-1998
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Superko, Colleen
 REGISTRATION NUMBER: 39,850
 REFERENCE/DOCKET NUMBER: 106.941.155
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 526-6000
 TELEFAX: (617) 526-5000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 619 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-066-046-2
 Query Match 6.0%; Score 112.5; DB 3; Length 619;
 Best Local Similarity 20.2%; Pred. No. 0.065;
 Matches 81; Conservative 74; Mismatches 158; Indels 87; Gaps 18;
 QY 12 IHALNTTKRAISTKNAIPILSS-----IKIEVSTGTGTLTSGNGQISIENTIPVSNENAG 66
 Db 215 VNQVNTT--AVQVSESLIPESQHTDVTLEDTTITVDGEYGHFSDIASGEHNDLPA 272
 QY 67 LLITSPGAILLEASFPFNIISLPDISI--NVKEIQHVLTSGKSEITLKGKVDQY 124
 Db 273 MLLDEADFTMLLANEESKLTLESMPDSLEDNVQSLG---TLPIQEGETVSEGNTRRESLP 328
 QY 125 RLOQVSTENPILKTLKLSIIAETAFASLQESRPLTGVHIVLSNKHDPKAVTD--S 182
 Db 329 --TVDSQDSVGVSTDLLEAHQSEVETVSEVTQDS--LSTNI-----SQDSVGVSTDL 378
 QY 183 HRMSQRLI---TLDNSTADLMVLPKSLRFSFAVFTDDIETVEVFSPS----- 229
 Db 379 HSKGVEIVSEGGTQDSLSADFPINTVSESTDLLEA-HSQEIVTSEFTQDSLSTNISQDS 437
 QY 230 -----QILPSEHISFYTRLLEGNYPTDRLLMTETETEVFTQSRLHAMERAFILSNA 284
 Db 438 VGVSTDLLEHVSQSEIVEIS---EGGTQDS---LSTNISQDSVGVSTDLLEAHQSEVETVSEF 491
 QY 285 TQNG-----TVKLEI-----TONHISAHVNSPVEGVKNEDL----- 315
 Db 492 TQDSLSTNISQDSVGVSTDLLEHVSQSEIVEIVSEGGTQDSLSTNISQDSVGV--VSTDLEH 550
 QY 316 --DIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSLP 353
 Db 551 GVEIVSEGGTQDSLSDAD-----FPINTVSESTDLLEHSP 585

RESULT 19
 US-09-248-796A-19154
 ; Sequence 19154, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 19154
LENGTH: 1753
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: UNSURE
LOCATION: (223)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-19154

Query Match
Best Local Similarity 5.9%; Score 109.5; DB 4; Length 1753;
Matches 91; Conservative 80; Mismatches 145; Indels 121; Gaps 25;
QY 7 NRTFLHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTG-----SNQI-----SIE 55
DB 977 NRTTILE-LSLNPPQINNKK---LTLDLKADPKQGVFLAVFNITNGVWVWSISPIIK 1031
QY 56 NTPVSNENAGLLITSPGAILLEASFFINIISLP---DISINVKEIE-OHQVLTSGKS 111
DB 1032 GTIPM-----LDLTEDGSI-----FQIDNKLPIGNAINVKVKQVDLQHOILVLTARK 1079
QY 112 EITLAGKVDQ---YP-RLOEVSTENPLILTKLKLSIIAETAFASLOESR---PILTVGHVLSNHK--- 173
DB 1080 NFEIKFEDVRQQTTPARIKIKPNHVLV---ELGNVVIASSFVTDALNDYSDKLDHVP- 1135
QY 168 VLSNHKDKAVATDSHRMSQRLITLDNTSADLMVVL-----PSKSLREFSAV----- 214
DB 1136 ---NVNDY-----VSAKVLDDIDSTNKRISVSLRTDKAGSKDVINSISDLTRGVQIK 1183
QY 215 -PTDIIETVEVFPFSQILP-----RSEHISFVTRLEGNYP-----DTRLL 256
DB 1184 GFVKNISNGVYVSLGRSIYALVRVSDLSYLDKQWQNFKNQPVIGKIVNCKQEGRIL 1243
QY 257 MTEFETEV-----VENTQSLRHAMEAFILSNATQNGT-VKLEITONHISAHVNSPE 307
DB 1244 MTLKSEVNGELKIMKTFDDQLQIDFECT--VTSTDFGVFKLDGTVN-VSGLCHHSE 1300
QY 308 VGKVNED-----LDIVSQSGDLTISFNPTYLI--ESLKAISKSETVKIH 349
DB 1301 ISENDVDNVKLLFGTCDRVKVKILKIDSEK-RQLSLGMKASFYITDSKEDNDDDEME 1359
QY 350 FLSPVRPFTLTGDEEE 366
DB 1360 DASEV-----GEEEEE 1369

RESULT 20
US-09-710-279-546
Sequence 546, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 546
LENGTH: 532
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-546

Query Match
Best Local Similarity 5.9%; Score 109; DB 4; Length 532;
Matches 91; Conservative 80; Mismatches 145; Indels 121; Gaps 25;

Matches 76; Conservative 76; Mismatches 160; Indels 78; Gaps 15;
QY 19 KRAI-----STKNAIPILSS-----IKIEVTSTGVTLTGNGQ-ISISNTI 58
DB 34 KKAIIIEGIFDIDESKDAINILSLAIDVDEDFLLVKREIFFSSGKSIICRINNQTIVLQDLR 93
QY 59 PVSNEAGLLITSPGAILLEASFFINIISLPDISINVKEIEHQVLTSGKSEITLKGK 118
DB 94 KVMQELLDIHGHQETQSLKOKYHLQLLDYADNOYS-DLLNQYQLSYNQYKN-----RKK 148
QY 119 DVDOYPRLOEVSTENPLILTKLKLSIIAETAFASLOESR---PILTVGHVLSNHK--- 173
DB 149 ELEELE-----SAQALLQRLDLMKFQLEET-EASLKEGEVDQLESIDKRONSEKLN 202
QY 174 ---DFKAVATDSHRMSQRLITLDNTSADLMVVLPSKSLREFSAVFTDIIETVEVFPSPQ 230
DB 203 ALNNAHQVLTDESAPDRLYELSNVLOQTINDIVPEKFR-----LKEDID----- 247
QY 231 ILPRSEHISFVTRLEGNYPDTDRLLMTPEFTEVVFNTQSLRHAMEAFILSNATQNGTV 290
DB 248 -----QFYMLLEDKAKHEIYDEKANTFDEQVNLNYES-----RMNLLNLRK--Y 291
QY 291 KLEITONHISAHVNSPEVGVKNEDLDIVSQSGDLTISFNPTYLIETSLKAISKSETVKI-- 348
DB 292 GKDIETELAYOSKLANEIDKIEYEQSTQSLREBIKTYNE--VIDICKLSQERRRVAR 349
QY 349 ---HFLSPVRPFTLTGDEEESFQLITP 374
DB 350 ELRDHIVSEIQNLQMKDANLEISPKPLDEP 379

RESULT 21
US-09-248-796A-16697
Sequence 16697, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 16697
LENGTH: 569
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-16697

Query Match
Best Local Similarity 5.8%; Score 107.5; DB 4; Length 569;
Matches 80; Conservative 56; Mismatches 171; Indels 65; Gaps 13;
QY 14 ALNTTKRAISTKNAIPILSSIKIEVTST--GVTLTSGNGQISIENTIPVSNENAGLLITS 71
DB 178 ASSSTKSSVSVDLT--VVNSILLSETSLSSDDTSSTDSISSTTNSDTGNINAGSLHTS 235
QY 72 PGAI-----LLEASFFINIISLPDISINVKEIEHQVLT----- 107
DB 236 TASIKELSIQKTGVTLSLSSSYLSLSTKLSSTSDITIELITTELTITIEDNEPNTFTSP 295
QY 108 SGKSEI-----TLKGKVDQYPRLOEVSTENPLILTKLKLSIIAETAFASLOESRPL 162
DB 296 SSHSEIFSSDNLVSKQVDRESTIKTPTDVTTVSSLSVHSTEASTATLGNSFSNVA 355
QY 163 TGVHIVL-----SNHKDFKAVATDSHRMSQRLITLDNTSADLMVVLPSKSLREFSAV 214
DB 356 TPLNIATSLRSTSSSSNHAT-ESSCTVKSEASAEIIPSPPTSTDNRL-----SYSTEAKGI 411

Query Match	5.6%;	Score 104.5;	DB 4;	Length 1726;
Best Local Similarity	18.2%;	Pred. No. 2;		
Matches 72;	Conservative 61;	Mismatches 147;	Indels 115;	Gaps 13;

Db 502 VLQDIKVDATSV-KLSDEAVGGEVE-----GWNQKTIKVSADVQYKEIKGTGK 552
Qy 135 LILKTK-----LLKSIATAFAASLOESRPILTGVHIVLSNH-----KDFKAVAT 180
Db 553 VVTVTNTEGLVKNVNSDNTIDF-----DSGNSATDQFVVVATKDKIVNGKVEVKFKN-AS 608
Qy 181 DSHRMSQRLITL-----DNTSADLMVVLPSKSLRFSAVFTDDIETVEVFFSPSILF 233
Db 609 DTTPTSTKTIIVNVVNVKADATPVGLDIVARS-----EIDVNPNTASTADVDF-----I 658
Qy 234 RSEHISFVTRLEGN-----YPTDRLLMTFETEVFNTOQSLRHAMERAFILSNATQNG 288
Db 659 NFESVEIYTLDSNGNRLKVKVTPATTLVGTNDYVEVNGVNLQFKGNDELTL-----TSSS 714
Qy 289 TVKLEITONHISAHV-----NSPEVKVKNEDLDIVSQSGDLTIS----- 328
Db 715 TVNVDVTADGITKRIPIVKVINSASVPASATVATSPVTVKLN-----SSNDLTFEELI 767
Qy 329 ---FNPTVLIIE 336
Db 768 FGVIDPTQLVK 778

RESULT 27

US-09-889-572-4
; Sequence 4, Application US/09889572
; Patent No. 6610517
; GENERAL INFORMATION:

; APPLICANT: Lubitz, Werner
; TITLE OF INVENTION: COMPARTMENTALIZATION OF RECOMBINANT POLYPEPTIDES IN HOST CELLS
; FILE REFERENCE: 100564-00070
; CURRENT APPLICATION NUMBER: US/09/889,572
; PRIORITY FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: PCT/EP00/00686
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: DE 199 03 345.5
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-889-572-4

Query Match 5.6%; Score 103.5; DB 4; Length 921;
Best Local Similarity 20.8%; Pred. No. 0.91;
Matches 77; Conservative 59; Mismatches 146; Indels 89; Gaps 16;

Qy 16 NTTKRAISTKNAIPILSSIKIEVTSTGYTL-TGSGNGQISIENTIPVSNENAGLLITSPGA 74
Db 448 DATVRSLN-----PIATAANGSELLVTANAGGKASFEVTFKDNKRTFTVDVKKDP 502
Qy 75 ILLEASFFINISSLIPDISINVKIEQHGVVLTSKSEITLKGKDVQDQYPRLOEVSTENP 134
Db 503 VLQDIKVDATSV-KLSDEAVGGEVE-----GWNQKTIKVSADVQYKEIKGTGK 553
Qy 135 LILKTK-----LLKSIATAFAASLOESRPILTGVHIVLSNH-----KDFKAVAT 180
Db 554 VVTVTNTEGLVKNVNSDNTIDF-----DSGNSATDQFVVVATKDKIVNGKVEVKFKN-AS 609
Qy 181 DSHRMSQRLITL-----DNTSADLMVVLPSKSLRFSAVFTDDIETVEVFFSPSILF 233
Db 610 DTTPTSTKTIIVNVVNVKADATPVGLDIVARS-----EIDVNPNTASTADVDF-----I 659
Qy 234 RSEHISFVTRLEGN-----YPTDRLLMTFETEVFNTOQSLRHAMERAFILSNATQNG 288
Db 660 NFESVEIYTLDSNGNRLKVKVTPATTLVGTNDYVEVNGVNLQFKGNDELTL-----TSSS 715
Qy 289 TVKLEITONHISAHV-----NSPEVKVKNEDLDIVSQSGDLTIS----- 328
Db 716 TVNVDVTADGITKRIPIVKVINSASVPASATVATSPVTVKLN-----SSNDLTFEELI 768

Qy 329 ---FNPTVLIIE 336
Db 769 FGVIDPTQLVK 779

RESULT 28

US-09-134-000C-4971
; Sequence 4971, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4971
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4971

Query Match 5.5%; Score 103; DB 4; Length 1233;
Best Local Similarity 20.5%; Pred. No. 1.6;
Matches 86; Conservative 75; Mismatches 115; Indels 144; Gaps 24;

Qy 46 TGSNGQI--STENTIPVSNENAGLLIT-SPGAIL-----LEASFFINISSLIPDISINVK 97
Db 263 TGLDGEIFYNDRTL-TGNQTLLELTTPGAVFGKQDNLEPQVFSY-----DVDINGQ 315
Qy 98 EIEQHGVVLTSKSEITLKGKDVQDQYPRLOEVSTENPILKTLKLSIAETAFAASLOE 157
Db 316 ILPETQTLTTPGK-DYTL-----SDNSLGRITAVTPVNMNQKAYSLSINR 359
Qy 158 SRPILTGHIIVLSNHKDFKAVATDSHRMSQRLITLNTSADLMVVLPSKSLRFSAVTD 217
Db 360 T-----IYLESADY-----NYLYSQOYPTTKIGSISLKSSTGTGKTTDTAKTSQ 405
Qy 218 DIETV--EVFPSPQIILFRSE--HISFYTRLLRGVNPDTDRLLMTFETEVWF---NTQ 269
Db 406 TSKVIADREMSYISFQSGKYVIYITLSE-----TKVGOQIVLESTNGQ 454
Qy 270 SLRHAMERAF-----LISNATQNGTVKLEIT-----QNHISA 301
Db 455 EIKNPKFTAYGPLYENVKLEDYFDIKTEGGKLTLTATKDSYLINISDLTMDFKKDINL 514
Qy 302 HVNSPEVKVKNEDLDIVS-----QSG---SDLTISP----- 329
Db 515 SLSTPVTIGP-NKAIQLVSDQYIEPISVNVPLNAETAMGNYDONGAYSRTTIVSNGSKEK 573
Qy 330 -----NPTVLESKAIKSETVKIHFSPV-RPFTLTGDEESESFTQLITPVRTN 378
Db 574 PIONLEIKVHPNVL--SLRATK-----EIVFYKLGTDYVTP-TSDGSVIKFTTPI-TN 625

RESULT 29

US-09-071-035-404
; Sequence 404, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:

; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland

COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 404:
SEQUENCE CHARACTERISTICS:
LENGTH: 1416 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-404

Query Match 5.5%; Score 103; DB 4; Length 1416;
Best Local Similarity 20.5%; Pred. No. 2.1;
Matches 86; Conservative 75; Mismatches 115; Indels 144; Gaps 24;
QY 46 TGSNGQI--SIENTIPVSNENAGLLIT-SPGAIL-----LEASFFINISSLPDISINVK 97
DB 232 TGLDGEIYNLDRTL-TGNQLELLTTPGAVFGKQDNLEPQVFSY-----DVDINGQ 284
QY 98 EIEQHVLTGSKSEITLKGKDVQYPRLQEVSTENPLILKTLKLSIIAETAFASLOE 157
DB 285 ILPQTQTLTPGK-DYTL-----SDNSLGRIAVTPVNNMQKAYSLSINR 328
QY 158 SRPILTGVIHVLNHNKDFKAVATDSHRMSQRLITLDNTSADLMVVLPSKSLREFSAVTD 217
DB 329 T-----IYLESASDY-----NYLSQOQYPTTKIGSISLKSTGKTQTTDFTAQTSQ 374
QY 218 DIETV--EVFPSPQILFRSE---HISFYTRLLEGNYPTDRLLMTETETEVVF---NTQ 269
DB 375 TSKVIADREMRMSYISFQSKGYVYTYGLTE-----TKVQQQIVLESTNGQ 423
QY 270 SLRHAMERAF-----LISNATONGTVKLEIT-----QNHISA 301
DB 424 EIKPKFTAYGLYENVKLEDYFDIKTEGGKLTATKDSYLRLINISDLTMDFKKDINL 483
QY 302 HVNSPEVGNEDDIVS-----QSG---SDLTISF----- 329
DB 484 SLSTPVGIP-NKAIQLVSDQYIEPISVVNPNAETAWGNYDQNGAYSRRITVSVNGSKEK 542
QY 330 -----NPTYLIESLKAISKETVKIHFLSPV-RPFTLTPGDEESFIQLITPVRTN 378
DB 543 PQNLKIKVGHENYL--SLRATK-----EIFYFYKLGTDYVTP-TSDGSVIKFTTPI-TN 594

RESULT 30
US-09-071-035-402
Sequence 402, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 402:
SEQUENCE CHARACTERISTICS:
LENGTH: 1448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-402

Query Match 5.5%; Score 103; DB 4; Length 1448;
Best Local Similarity 20.5%; Pred. No. 2.1;
Matches 86; Conservative 75; Mismatches 115; Indels 144; Gaps 24;
QY 46 TGSNGQI--SIENTIPVSNENAGLLIT-SPGAIL-----LEASFFINISSLPDISINVK 97
DB 259 TGLDGEIYNLDRTL-TGNQLELLTTPGAVFGKQDNLEPQVFSY-----DVDINGQ 311
QY 98 EIEQHVLTGSKSEITLKGKDVQYPRLQEVSTENPLILKTLKLSIIAETAFASLOE 157
DB 312 ILPQTQTLTPGK-DYTL-----SDNSLGRIAVTPVNNMQKAYSLSINR 355
QY 158 SRPILTGVIHVLNHNKDFKAVATDSHRMSQRLITLDNTSADLMVVLPSKSLREFSAVTD 217
DB 356 T-----IYLESASDY-----NYLSQOQYPTTKIGSISLKSTGKTQTTDFTAQTSQ 401
QY 218 DIETV--EVFPSPQILFRSE---HISFYTRLLEGNYPTDRLLMTETETEVVF---NTQ 269
DB 402 TSKVIADREMRMSYISFQSKGYVYTYGLTE-----TKVQQQIVLESTNGQ 450
QY 270 SLRHAMERAF-----LISNATONGTVKLEIT-----QNHISA 301
DB 451 EIKPKFTAYGLYENVKLEDYFDIKTEGGKLTATKDSYLRLINISDLTMDFKKDINL 510
QY 302 HVNSPEVGNEDDIVS-----QSG---SDLTISF----- 329
DB 511 SLSTPVGIP-NKAIQLVSDQYIEPISVVNPNAETAWGNYDQNGAYSRRITVSVNGSKEK 569
QY 330 -----NPTYLIESLKAISKETVKIHFLSPV-RPFTLTPGDEESFIQLITPVRTN 378
DB 570 PQNLKIKVGHENYL--SLRATK-----EIFYFYKLGTDYVTP-TSDGSVIKFTTPI-TN 621

RESULT 31
US-09-134-001C-5119
Sequence 519, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: GTC-007

;; CURRENT APPLICATION NUMBER: US/09/134,001C

;; CURRENT FILING DATE: 1998-08-13

;; PRIOR APPLICATION NUMBER: US 60/064,964

;; PRIOR FILING DATE: 1997-11-08

;; PRIOR APPLICATION NUMBER: US 60/055,779

;; PRIOR FILING DATE: 1997-08-14

;; NUMBER OF SEQ ID NOS: 5674

;; SEQ ID NO 5119

;; LENGTH: 395

;; TYPE: PRT

;; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5119

Query Match

Best Local Similarity 5.5%; Score 102.5; DB 3; Length 395;

Matches 79; Conservative 59; Mismatches 135; Indels 117; Gaps 17;

QY 30 ILSSIKIEVTSTGVTLTGSGNGQISIENTIPV-SNENAGLITSPCAILLASFFINI1SS 88

Db 28 ILSAQLPMVGHRSITDFEIASEAPKGLKPVFGSKNEVLILTSSTSLASM----- 80

QY 89 LPDISINVKIEBQHVLTSGKSEITLK--GKVDQVPRLOEVSTENFLILK---TKLLK 143

Db 81 -----LNIANPDHIVIIVSGAFGNRFKQIAQTYNHVHVYDNNWGEAIVVDVFTYLLKQ 135

QY 144 SIIAETAFASLOESRPILTV-HIVLSNHKDFKAVATDSHRMSORLITLNT----- 195

Db 136 LNVPTAVFTQCETS---TGVHFPV-----HQLGHAKAFDNLSTFYVDG 178

QY 196 -----SADLMVLPKSLREFS-----AVFTD-----DIETVEVF----- 225

Db 179 VSCIGAVDVLTKIDIVLVSQKAIMLPGLAFVAVSDRAKKFADVKTPRFYDLNKK 238

QY 226 -----FSPQILFRSEHISFYTRLLEGNYPTDRLMLTEFETEVVNTQSLRHA 274

Db 239 YIKSQEQNSTPTPTNVGLFRG--INAYVELVK-----KEGLNHV 275

QY 275 MERAFLISNATQGTVKLEI-----TQNHSAHVNSPVGVKNEDDIVSQSGDLTISF 329

Db 276 ISRHFKIRNALRAALKALELLVKDDAHAPTPTS-FVPRKQBELNLIK---NOLKSQF 331

QY 330 NPTYLIESLKAIKSETVKHFLSPVRFTL 359

Db 332 NIT-IAGQGHKLGQILRAIGHMKISPFDI 360

RESULT 32

US-09-248-796A-23039

;; Sequence 23039, Application US/09248796A

;; Patent No. 6747137

;; GENERAL INFORMATION:

;; APPLICANT: Keith Weinstock et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

;; FILE REFERENCE: 107196.132

;; CURRENT APPLICATION NUMBER: US/09/248,796A

;; PRIOR FILING DATE: 1999-02-12

;; PRIOR APPLICATION NUMBER: US 60/074,725

;; PRIOR FILING DATE: 1998-02-13

;; PRIOR APPLICATION NUMBER: US 60/096,409

;; PRIOR FILING DATE: 1998-08-13

;; NUMBER OF SEQ ID NOS: 28208

;; SEQ ID NO 23039

;; LENGTH: 461

;; TYPE: PRT

;; ORGANISM: Candida albicans

US-09-248-796A-23039

Query Match

Best Local Similarity 5.5%; Score 102.5; DB 4; Length 461;

Matches 64; Conservative 76; Mismatches 154; Indels 51; Gaps 11;

QY 19 KRAISTKNAIPILSIIKIEVTSTGVTLTGSGNGQISIENTIPVSNENAGLITSPCAI--L 76

Db 5 KSVLPVKGALPVKGPALPSEVSSSS--EVSSSQVISSEVSSSSSE---VVSSSEVSSSS 59

QY 77 LEASFFINISSLPDIS-----INVKEIEHQHVLTSGKSEITLKGD 119

Db 60 SEVSSSEVSSSSSEVSSSSSEVSSSSSQVTSSEIIVSSSEVSSSEVSSSEVSSSSSE 119

QY 120 VDOYPRL---QEVSTENPLIKTKLLKSIIAETAFASLOESRPILTVGHVILSNHK-- 173

Db 120 VSSSEVSSSSSEVSSSSSEVSSSSSQVTSSEIIVSSSEVSSSSSQVVISSEVSSSEVS 177

QY 174 -DFKAVATDSH-RMSQRLITLNTSADLMVLPKSLREFAVFTDDITV----- 222

Db 178 SSSEVSSSSSEVSSSSSEVSSSSSQVTSSEIIVSSSEVSSSSSEVSSSSSEVSSSS 237

QY 223 -EVFFSPQILFRSE---HISFYTRLLEGNYPTDRLMLTEFETEVVNTQSLRHAWE 276

Db 238 SEVSSSEVSSSSSEVSSSSSQVTSSEIIVSSSEVSSSSSEVSSSSSEVSSSSSE 297

QY 277 RAFLISNATQGTVKLEITQNHSAHVNSPVGVKNEDDIVSQS 321

Db 298 VVSSSEVSSSSSEVSSSSSQVTSSEIIVSSSEVSSSSSEVSSSSSEVSSSSSE 340

RESULT 33

US-08-480-604A-10

;; Sequence 10, Application US/08480604A

;; Patent No. 5736139

;; GENERAL INFORMATION:

;; APPLICANT: KINK JOHN A.

;; APPLICANT: THALLEY, BRUCE S.

;; APPLICANT: PADHYE, NISHA V.

;; APPLICANT: FIRCA, JOSEPH R.

;; APPLICANT: STAFFORD, DOUGLAS C.

;; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND

;; PREVENTION OF C. DIFFICILE DISEASE

;; NUMBER OF SEQUENCES: 32

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: MEDLEN & CARROLL, LLP

;; STREET: 220 MONTGOMERY STREET, SUITE 2200

;; CITY: SAN FRANCISCO

;; STATE: CALIFORNIA

;; COUNTRY: UNITED STATES OF AMERICA

;; ZIP: 94104

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/480,604A

;; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/422,711

;; FILING DATE: 14-APR-1995

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/405,496

;; FILING DATE: 16-MAR-1995

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/329,154

;; FILING DATE: 25-OCT-1994

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/161,907

;; FILING DATE: 02-DEC-1993

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/985,321

;; FILING DATE: 04-DEC-1992

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/429,791

;; FILING DATE: 31-OCT-1989

```

; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40, 027
; REFERENCE/DOCKET NUMBER: OPHD-01763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-604A-10

Query Match 5.5%; Score 102; DB 1; Length 2366;
Best Local Similarity 20.9%; Pred. No. 6;
Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;

Qy 30 ILSSIKI-----EVSTGTVLTCGNGQISIENTIPVSNENAGLLITSPAILLEA 79
Db 1371 ILSTLSTEENKILNHSHEINPFG-EVNGSGFVSLTFSI-LEGINAIEVD-----LLSK 1423
Qy 80 SFFNIILSLPDISINVKETEIOHVLTSGKSEITLKGKVDQYPRLOEVSTENPLI--- 136
Db 1424 SYKLLISGEIKILMLNSHNHQQ-KIDYIGFNSELQ---KNI-PYSFVDSGKNGFNGS 1478
Qy 137 LKTKLLKSIITAETAFAAS--LQESRPILTGVHIVLSNHKDFKAVATDSHRMSQRLITLDN 194
Db 1479 TKEGLFVSELPDVLVLSKVMDSDKP---SFGYYNNLKVVKVITKONVILTYLKKDD 1535
Qy 195 TSADLMVLPKSLRPFSAVTDIDTVEVFFSPSQILPRSEHISFYTRLLEGNYPDTR 254
Db 1536 IKISLSLTQDEKTIKLSNVHLDESGVAEI-----LKFWR--KGNWNTSDS 1580
Qy 255 LL-----MTFEFEVVTQSLRHAMERAFNISNATONGT-----VKLEI 294
Db 1581 LMSFLESMNKISIPWNLQSNIRKILFDANFIISGTTISIGOFETCDENDNIQPYFIKFT 1640
Qy 295 TQNHISAHVNSPE--VGKVNEDLDIVSQSG--SDLTISFNPTYL 334
Db 1641 LETNYTLVGNRQMVIEPNVYLLD---DSGDISVTINFSQKYL 1681

RESULT 34
US-08-405-496A-10
; Sequence 10, Application US/08405496A
; Patent No. 5919665
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, JAMES A.
; TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
; TITLE OF INVENTION: NEUROTOXIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-405-496A-10

Query Match 5.5%; Score 102; DB 2; Length 2366;
Best Local Similarity 20.9%; Pred. No. 6;
Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;

QY 30 ILSSIKI-----EVTSTGVTLTGNGQISIENTIPVSNENAGLLITSPCAILLEA 79
DB 1371 ILSTLSIEENKIILNSHEINFSG-EVNGSNGFVSLTFSI-LEGINAIEVD-----LLSK 1423
QY 80 SFINIISLPDISINVKEIQHVLTSGKSEITLKGKDQVDPRLQEVSTENPLI---136
DB 1424 SYKLLISGELKILWANSNHIQQ-KIDYIGFNSQL---KNI-PYSFVDSGKENGFIN 1478
QY 137 LKTKLKLSIIAETAPFAS--LQSRPILTVHVLNHNKDFKAVATDSHRMSORLITLDN 194
DB 1479 TKEGLFVSELPDVVLISKVYMDSDKP---SFGYYSNNLKDKVKVITKDNVNILTYLKD 1535
QY 195 TSADLMVVLPSKLSREFSAVFTDIDIEVEVFPSPQILFRSEHISFVTRLEGNYPDTR 254
DB 1536 IKISLSTLQDEKTIKLNSVHLDSGVAEI-----LKFMR--KGNVTNTSDS 1580
QY 255 LL-----MTEFEVTVFNTQSLRHAMERAPLISNATQNGT-----VKLEI 294
DB 1581 LMSFLESMNIKSIFVNFLOSNIKFILDANFIISGTTISIGOFFEICDENDNIQYFIKNT 1640
QY 295 TQNHSAHVNSPE--VGKVNEDLDIVSQSG--SDLTISFNPTYL 334
DB 1641 LETNYTLVGNRQNMIVEPNYDLD---DSGDISSTVINFSQKYL 1681

RESULT 35
US-08-915-136-10
; Sequence 10, Application US/08915136
; Patent No. 6290960
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,136
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/480,604
; FILING DATE:
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPD-01763
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-915-136-10

```

```

Query Match      5.5%; Score 102; DB 3; Length 2366;
Best Local Similarity 20.9%; Pred. No. 6;
Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;

QY 30 ILSSIKI-----EVTSTGVTLTGSGQISIENTIPVSNENAGLLITSPGAILLEA 79
Db 1371 ILSTLSIENKIILNSHINFSG-EVNGSGFVSLTFSI-LEGINAIIIEVD-----LLSK 1423

QY 80 SFFNIISLPDISINVKIEHQVLTSGKSEITLKGKVDQYPRLOEVSTENPLI---136
Db 1424 SYKLLISGELKILMLNSHIIQ-KIDYIGFNSLQ---KNI-PYSFVDSGKENGFINGS 1478

QY 137 LKTKLLKSIITAFAAS--LOESRPILTGVIHVLNSHKDFKAVATDSHRMSQRLITLDN 194
Db 1479 TKEGLFVSELPDVVLISKVYMDSDKP---SFGYYSNNLKDKVITKDNVILTYGKLD 1535

QY 195 TSADLMVLPKSLRFSFAVTDDETVEVFFSPQILFRSEHISFYTRLLGPNPTDR 254
Db 1536 IKISLSLTLDQEKTKLNSVHLDSEGVAEI-----LKFMR--KGNNTSDS 1580

QY 255 LL-----MTEFETEVNTQSLRHAMERAFILSNATQNGT-----VKLEI 294
Db 1581 LMSFLESNNIKSIFVNFLOQNIKFILDANFIISGTTSGQEFICDENDNIQPYFIKENT 1640

QY 295 TQNHISAHVNSPE--VGKVNEDLDIVSOG--SDLTISFNPTYL 334
Db 1641 LETNYTLVGNRQNVIFENYDLD---DSGLISSTVINFSOKYL 1681

```

RESULT 36

US-08-957-310-10

; Sequence 10, Application US/08957310

```

; Patent No. 6365158
; GENERAL INFORMATION:
; APPLICANT: Williams, James A.
; APPLICANT: Kink, John A.
; TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
; TITLE OF INVENTION: OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,310
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 24-OCT-1994
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPD-01121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-957-310-10

Query Match      5.5%; Score 102; DB 3; Length 2366;
Best Local Similarity 20.9%; Pred. No. 6;
Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;

QY 30 ILSSIKI-----EVTSTGVTLTGSGQISIENTIPVSNENAGLLITSPGAILLEA 79
Db 1371 ILSTLSIENKIILNSHINFSG-EVNGSGFVSLTFSI-LEGINAIIIEVD-----LLSK 1423

QY 80 SFFNIISLPDISINVKIEHQVLTSGKSEITLKGKVDQYPRLOEVSTENPLI---136
Db 1424 SYKLLISGELKILMLNSHIIQ-KIDYIGFNSLQ---KNI-PYSFVDSGKENGFINGS 1478

QY 137 LKTKLLKSIITAFAAS--LOESRPILTGVIHVLNSHKDFKAVATDSHRMSQRLITLDN 194
Db 1479 TKEGLFVSELPDVVLISKVYMDSDKP---SFGYYSNNLKDKVITKDNVILTYGKLD 1535

QY 195 TSADLMVLPKSLRFSFAVTDDETVEVFFSPQILFRSEHISFYTRLLGPNPTDR 254
Db 1536 IKISLSLTLDQEKTKLNSVHLDSEGVAEI-----LKFMR--KGNNTSDS 1580

QY 255 LL-----MTEFETEVNTQSLRHAMERAFILSNATQNGT-----VKLEI 294
Db 1581 LMSFLESNNIKSIFVNFLOQNIKFILDANFIISGTTSGQEFICDENDNIQPYFIKENT 1640

```

QY 295 TQNHSAHVNSPE--VGKVNEDLDIVSQSG--SDLTISFNPTYL 334
 Db 1641 LETNYTLVGNRQNMIVEPNYDLD---DSGDISSTVINFQSKYL 1681

RESULT 37

US-10-011-366-10
 ; Sequence 10, Application US/10011366
 ; Patent No. 6573003
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, James A.
 ; TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
 ; OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
 ; DISEASE

NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Medlen & Carroll
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/10/011,366
 FILING DATE: 16-NOV-1993
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/957,310
 FILING DATE: 23-OCT-1997
 APPLICATION NUMBER: US/08/329,154
 FILING DATE: 24-OCT-1994
 APPLICATION NUMBER: US/08/161,907
 FILING DATE: 02-DEC-1993
 APPLICATION NUMBER: US/07/985,321
 FILING DATE: 04-DEC-1992
 APPLICATION NUMBER: US/07/429,791
 FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:
 NAME: Ingolia, Diane E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01121
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2366 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-10-011-366-10

Query Match 5.5%; Score 102; DB 4; Length 2366;
 Best Local Similarity 20.9%; Pred. No. 6;
 Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;

QY 30 ILSSIKI-----EVTSTGTLTSGNQISIENTIPVSNENAGLITSPGAILLEA 79
 Db 1371 ILSTLSIEENKIILNSHIEINFG--EYNGSNGFVSLTFSI--LEGINAIIEVD-----LLSK 1423
 QY 80 SFFINISSLPDISINVKEHQHVLTSGSKSITLTKGKDQVQYPRLQVSTENPLI--- 136
 Db 1424 SYKLLISGELKILMNSNIHQ--KIDYIGFNSBLQ---KNI--PYSFVDSGKNGFINGS 1478
 QY 137 LKTKLLKSIATAFAAS--LOESRPILTGVHVLNSHKDFKAVATDSHRMSQRLITLON 194

Db 1479 TKEGLFVSELDPVLIISKVYMDSKP---SFGYYNNLKDVKVITKDNVNIITGYLKKDD 1535
 QY 195 TSADLMVVLPSKSLREFSAVFTDDTETVEVFSPQILFRSEHISFYTLLEGNYPDTR 254
 Db 1536 IKISLSLTLODEKTIKLSVHLDESQVAEI-----LKFMNR--KGNTNTSDS 1580
 QY 255 LL-----MTEFETEVFVNTQSLRHAMERAFILSNATQNGT-----VKLEI 294
 Db 1581 LMSFLESMNIKISIFVNFQSNIKFILDANFIISGTTISGQFEFICDENDNIQPYFIKFT 1640
 QY 295 TQNHSAHVNSPE--VGKVNEDLDIVSQSG--SDLTISFNPTYL 334
 Db 1641 LETNYTLVGNRQNMIVEPNYDLD---DSGDISSTVINFQSKYL 1681

RESULT 38

US-09-084-517-10
 ; Sequence 10, Application US/09084517
 ; Patent No. 6613329
 ; GENERAL INFORMATION:
 ; APPLICANT: KINK, JOHN A.
 ; APPLICANT: WILLIAMS, JAMES A.
 ; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
 ; PREVENTION OF C. DIFFICILE DISEASE
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
 ; STREET: 220 MONTGOMERY STREET, SUITE 2200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/084,517
 ; FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: CARROLL, PETER G.
 REGISTRATION NUMBER: 32,837
 REFERENCE/DOCKET NUMBER: OPHD-01610
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2366 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-084-517-10

Query Match 5.5%; Score 102; DB 4; Length 2366;

Best Local Similarity 20.9%; Pred. No. 6;
Matches 72; Conservative 61; Mismatches 139; Indels 72; Gaps 17;
QY 30 ILSSIKI-----EVTSGVLTGSGNGOISIENTIPVSNAGLLITSPGAILLEA 79
Db 1371 ILUSTIENKILMSHNEIFSG-EVNGSGVLTGSGNGOISIENTIPVSNAGLLITSPGAILLEA 79
QY 80 SFFINIISLPDISINVKIEHQVLTGSGSEITLKGKDVDPQYPRLOEVSTENPLI--- 136
Db 1424 SYKLLISGELKILMSNHIQ--KIDYIGFNSLQ---KNI-PYSFVDSGKNGFINGS 1478
QY 137 LTKLLKSIITATAFAS--LOESPILTGVIHVLNKHDFKAVATDSHRMSQRLITDN 194
Db 1479 TKEGFLSELPLVILISKYMDSKP---SFGYYSNNLKDVKVITKDNVNILTYGLKDD 1535
QY 195 TSADLMVPLPSKSLREFSAVETDIEVEVFPSPQILFRSEHISFYTRLLEGNYPDTR 254
Db 1536 IKISLSLTQDEKTKILMSVHLDGSAEI-----LKPMMR--KGNTNTSDS 1580
QY 255 LL-----MTEFETEVFNTQSLRHAMERAFLLISNATONGT-----VKLEI 294
Db 1581 LMSPLSMNIIKIFVFLQSNKIFLDANFLISGTTISGQFEFICDENDNIQPIKENT 1640
QY 295 TQNHISAHVNSE--VGKNEDELIVSQG---SDLTISFNPTYL 334
Db 1641 LETNYTLVGNRQNMIVENPYDLD--DSGDISSTVINFSQKYL 1681

RESULT 39
US-09-538-092-326
; Sequence 326, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 326
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number YGL238W
US-09-538-092-326

Query Match 5.4%; Score 101.5; DB 4; Length 960;
Best Local Similarity 24.4%; Pred. No. 1.5;
Matches 65; Conservative 44; Mismatches 94; Indels 63; Gaps 14;
QY 101 QHVVLTGSKSI-----TLKGKDVDPQYPRLOEVSTENPLIKTKLSIIATAFAA 153
Db 254 EHASVLIKVKSIIQELVQLVYTRYEDVFG-PMINEF-----IQITWNLLTISNQPKYDI 307
QY 154 SIQESRPILTGVHIVLSNKHDFKAVATDSHRMSQRLITDNADLMVPLPSKSLREFSA 213
Db 308 LVSKLSFLTAVTRI---PKYFEIENNES-----AMNITE--QIILPNVTLRE--- 351
QY 214 VFTDDIETVEVFPSPQILFRSEHISFYTRLLEGNYPDTRLLMTEF-----ETEWFVN 267
Db 352 ---EDVE-----LFEDDPIEYIRRDLEGGDTDRREACTDFLKEKEKNEVLVT 397
QY 268 TQSLRHAMERAFLL--ISNATONGTVK-----LEITQNHISAHVNSEPVGKVNEDLDIV-- 318

Db 398 NIFLAHM--KGFVDQYMSDPKSNKFKDLYLYLTALAINGNITNAGVSSNTNLLNVVDF 455
QY 319 --SQSGSLTISFNFNTYLIESLKAIAK 342
Db 456 FTKEIAPDLT-SNNIPHILRVDAIK 480
RESULT 40
US-09-543-681A-4674
; Sequence 4674, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4674
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4674

Query Match 5.4%; Score 101; DB 4; Length 385;
Best Local Similarity 20.3%; Pred. No. 0.38;
Matches 74; Conservative 58; Mismatches 123; Indels 110; Gaps 17;
QY 1 MIQFSINRTL-----FIHALNTKRAISTKNAIPILSSIKIEVTSITGV----- 43
Db 24 VINNKINETIGVSVTKPLYESVLOAKN-EKRSYARDLATYLVNSEAELTHARVGHDAKR 82
QY 44 -----TUTGNGQIS--IENTIPVS-----NENAGLLITSPGAILLEASF 81
Db 83 LHGDVRDILTALATVGEVKAITRNEIAVHEHLGEYTNARFNDHAGLIL-NPRA--LDLRF 139
QY 82 FINIISLPDISINVKIEHQV-----VLTSGKSEITLKGKDVDPQYPRLOEV 129
Db 140 FFWHWSIFALTEETARGIRHSIQFDDHGDLSLHKVYITDTHMDAWNLTIDTY-----L 194
QY 130 STENPL--ILKTKLLKSIITATAFAASIQESRPILTGVHIV-----LSNKHDFKAVA 179
Db 195 SPENPVLEITPAKSFDPVTTALAQQLEQQWRSMTDVHQPFIKILQENNLSPQAFKAVS 254
QY 180 TDSHRMSQRLITDNADLMVPLPSKSLREFSAVFTDDIETVEVFPSPQILFRSEHIS 239
Db 255 DD-----LAYQVONSALKTKLLAL-AKEVQNEIMIFVGNRGCVQIFTG----- 295
QY 240 FYTRLLEGNYPDTRLLMTEFETEV-----VNTQSLRHAME---RAPLISNATONGTV 290
Db 296 -----KIDRLVPHQFENSEQVWINIFNPAPFTLHLIESEIVESWITRKPTQDGFV 344
QY 291 -KLEI 294
Db 345 TSLEL 349

Search completed: January 28, 2005, 17:01:08
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2005, 16:47:03 ; Search time 71 Seconds
(without alignments)
3063.262 Million cell updates/sec

Title: US-10-048-071-28
Perfect score: 1863
Sequence: 1 MIQSFNRTLFIHALNTTKR.....LTPGDERSFIQLITPVRTN 378

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	378	Q9EVR1	Q9evr1 streptococc
2	1849	99.2	378	Q9A209	Q9a209 streptococc
3	1835	98.5	378	Q8P329	Q8p329 streptococc
4	1832	98.3	378	Q79YT3	Q79yt3 streptococc
5	1832	98.3	378	Q8K829	Q8k829 streptococc
6	1527	82.0	378	Q8DMN8	Q8dmn8 streptococc
7	1491	80.0	378	Q8E2I6	Q8e2i6 streptococc
8	1491	80.0	378	Q8E7Z3	Q8e7z3 streptococc
9	1389	74.6	378	DP3B_STRPN	O06672 streptococc
10	1384	74.3	378	DP3B_STRR6	P59651 streptococc
11	1053	56.5	380	DP3B_LACIA	O9cjl1 lactococcus
12	1052	56.5	380	DP3B_LACLC	O54376 lactococcus
13	316	45.2	189	Q9L571	Q9l571 streptococc
14	864.5	46.4	376	Q839Z4	Q839z4 enterococcu
15	830	44.6	379	Q890K7	Q890k7 lactobacilli
16	770.5	41.4	376	Q74M33	Q74m33 lactobacilli
17	770.5	41.4	376	AA507982	AA507982 lactobacili
18	748.5	40.2	377	DP3B_STAEP	Q8cqe6 staphylococ
19	730.5	39.2	377	DP3B_STAAM	P50029 staphylococ
20	730.5	39.2	377	Q6GD88	Q6gd88 staphylococ
21	730.5	39.2	377	Q6GKU3	Q6gku3 staphylococ
22	722.5	38.8	378	Q8EU87	Q8eu87 oceanobacil
23	710	38.1	381	Q92FV1	Q92fv1 listeria in
24	703	37.7	381	Q8YAW1	Q8yaw1 listeria mo
25	703	37.7	381	Q725G9	Q725g9 listeria mo
26	703	37.7	381	AA020792	AA020792 listeria
27	695	37.3	379	Q6HQ02	Q6hq02 bacillus th
28	693	37.2	379	Q73FK4	Q73fk4 bacillus ce
29	693	37.2	379	AA538938	AA538938 bacillus
30	692	37.1	379	Q81W34	Q81w34 bacillus an
31	692	37.1	379	AA029080	AA029080 bacillus

32	680.5	36.5	378	1	DP3B_BACSU	P05649 bacillus su
33	680	36.5	381	2	Q81JD4	Q81jd4 bacillus ce
34	679	36.4	381	2	Q8GQR3	Q8gqr3 bacillus we
35	624.5	33.5	380	1	DP3B_BACHD	Q9rcal bacillus th
36	569.5	30.6	376	2	Q6HI40	Q6hi40 bacillus th
37	557.5	29.9	376	2	Q81PM0	Q81pw0 bacillus an
38	557.5	29.9	376	2	AA031800	AA031800 bacillus
39	553.5	29.7	376	2	Q81CQ4	Q81cq4 bacillus ce
40	539.5	29.0	376	2	Q737D0	Q737d0 bacillus ce
41	539.5	29.0	376	2	AA541632	AA541632 bacillus
42	488.5	26.2	370	2	Q8RDL5	Q8rdl5 thermoanaer
43	457.5	24.6	366	2	Q97N34	Q97n34 clostridium
44	447	24.0	366	2	Q8XPG1	Q8xpg1 clostridium
45	433.5	23.3	378	2	Q6YRL2	Q6yrl2 onion yello

ALIGNMENTS

RESULT 1
Q9EVR1 PRELIMINARY; PRT; 378 AA.
AC Q9EVR1; (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 25, Last annotation update)
DE DNA polymerase III beta subunit (Fragment).
GN Name=dnaN;
OS Streptococcus pyogenes;
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Bruck I., O'Donnell M.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF280765; AAF98349.2; -;
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA_poliIII_beta.
DR Pfam; PF00712; DNA_poli3_beta_1.
DR Pfam; PF02767; DNA_poli3_beta_2; 1.
DR Pfam; PF02768; DNA_poli3_beta_3; 1.
DR SMART; SM00480; POL3BC; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
FT NON TER 378 378
SQ SEQUENCE 378 AA; 41866 MW; 4563F90B524C3D64 CRC64;

Query Match	100.0%	Score 1863;	DB 2;	Length 378;
Best Local Similarity	100.0%	Pred. No. 2.5e-115;	Mismatches 0;	Indels 0;
Matches 378;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MIQSFNRTLFIHALNTTKRAISTKNAIPILUSSKIEVTSTGVTLTGNGQISIENTIPV	60	
Db	1	MIQSFNRTLFIHALNTTKRAISTKNAIPILUSSKIEVTSTGVTLTGNGQISIENTIPV	60	
Qy	61	SNEAGLLITSPGAILLEASFPINTISLPDISINVKIEHQHVLTSGKSEITLKGKDV	120	
Db	61	SNEAGLLITSPGAILLEASFPINTISLPDISINVKIEHQHVLTSGKSEITLKGKDV	120	
Qy	121	DQYPRLOEVSTENPLILKLLKSLIIAETAFAASIQESRPILTGTVHIVLSNHKDFKAVAT	180	
Db	121	DQYPRLOEVSTENPLILKLLKSLIIAETAFAASIQESRPILTGTVHIVLSNHKDFKAVAT	180	
Qy	181	DSHRMSQRLITDNTSADLMVLPKSLRFSFAVFTDDIETVEVFFSPSQILFRSEHISF	240	
Db	181	DSHRMSQRLITDNTSADLMVLPKSLRFSFAVFTDDIETVEVFFSPSQILFRSEHISF	240	
Qy	241	YTRLLEGNYPTDRLMTTEFEVTVFNTQSLRHAMERAFILSNATQNTGVKLEITQNHIS	300	
Db	241	YTRLLEGNYPTDRLMTTEFEVTVFNTQSLRHAMERAFILSNATQNTGVKLEITQNHIS	300	

```
Qy 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVRPFTLT 360
Db 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVRPFTLT 360
Qy 361 PGDEESFIQLITPVRTN 378
Db 361 PGDEESFIQLITPVRTN 378

RESULT 2
Q9A209 PRELIMINARY; PRT; 378 AA.
AC Q9A209;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta subunit of DNA polymerase III (EC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=SPY0003;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAINS-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.C., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.B.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006472; AAK33147.1; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006260; P:DNA replication activity; IEA.
DR InterPro; IPR001001; DNA polIII beta.
DR Pfam; PF00712; DNA pol3 beta; 1.
DR Pfam; PF02767; DNA pol3 beta_2; 1.
DR Pfam; PF02768; DNA pol3 beta_3; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 378 AA; 41884 MW; 31DFF0FF326613D8 CRC64;

Query Match 99.2%; Score 1849; DB 2; Length 378;
Best Local Similarity 99.2%; Pred. No. 2.1e-114;
Matches 375; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Db 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Qy 61 SNEAGLLITSPGAILLEASFFINISSLDPDISINVKIEHQVVLTSKGEITLKGKDV 120
Db 61 SNEAGLLITSPGAILLEASFFINISSLDPDISINVKIEHQVVLTSKGEITLKGKDV 120
Qy 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGTHVILSNHKDPKAVAT 180
Db 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGTHVILSNHKDPKAVAT 180
Qy 181 DSHRMSQRLITLDNTSADLMVLPKSLRFEFAVFTDDIETVEVFFSPQILFRSEHISF 240
Db 181 DSHRMSQRLITLDNTSADLMVLPKSLRFEFAVFTDDIETVEVFFSPQILFRSEHISF 240
Qy 241 YTRLLEGNYPTDRLMLTETFEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLMLTETFEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
Qy 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVRPFTLT 360
```

```
Db 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVRPFTLT 360
Qy 361 PGDEESFIQLITPVRTN 378
Db 361 PGDEESFIQLITPVRTN 378

RESULT 3
Q8P329 PRELIMINARY; PRT; 378 AA.
AC Q8P329;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Beta subunit of DNA polymerase III.
GN Name=dnaN; OrderedLocusNames=spyM18_0002;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232; PubMed=11917108;
RX MEDLINE=21927593; PubMed=196838.1; -.
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Bares S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE009953; AAL96838.1; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA polIII beta.
DR Pfam; PF00712; DNA pol3 beta; 1.
DR Pfam; PF02767; DNA pol3 beta_2; 1.
DR Pfam; PF02768; DNA pol3 beta_3; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome.
SQ SEQUENCE 378 AA; 41863 MW; D2DB1C0CEFF16FD3 CRC64;

Query Match 98.5%; Score 1835; DB 2; Length 378;
Best Local Similarity 98.4%; Pred. No. 1.8e-113;
Matches 372; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Db 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Qy 61 SNEAGLLITSPGAILLEASFFINISSLDPDISINVKIEHQVVLTSKGEITLKGKDV 120
Db 61 SNEAGLLITSPGAILLEASFFINISSLDPDISINVKIEHQVVLTSKGEITLKGKDV 120
Qy 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGTHVILSNHKDPKAVAT 180
Db 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGTHVILSNHKDPKAVAT 180
Qy 181 DSHRMSQRLITLDNTSADLMVLPKSLRFEFAVFTDDIETVEVFFSPQILFRSEHISF 240
Db 181 DSHRMSQRLITLDNTSADLMVLPKSLRFEFAVFTDDIETVEVFFSPQILFRSEHISF 240
Qy 241 YTRLLEGNYPTDRLMLTETFEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLMLTETFEVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
Qy 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVRPFTLT 360
```

```

301 AHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFILSPVRPFTLT 360
361 PGDEESFIQLITPVRTN 378
361 PGDEESFIQLITPVRTN 378

RESULT 4
Q79YT3
ID Q79YT3 PRELIMINARY; PRT; 378 AA.
AC Q79YT3;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Putative DNA polymerase III beta subunit.
GN OrderedLocusNames=SPe0002;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_taxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1;
RX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
DR EMBL; AF005141; BAC63097.1; -
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
DR SEQUENCE 378 AA; 41861 MW; 24B19F5944A0E3B4 CRC64;

Query Match 98.3%; Score 1832; DB 2; Length 378;
Best Local Similarity 98.1%; Pred. No. 2.9e-113;
Matches 371; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MIQPSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Db 1 MIQPSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTPTGVTLTGNGQISIENTIPV 60
Qy 61 SNEAGLLITSPGAILLEASPFINISSLPDISINVKIEQHVLTSGKSEITLKGKDV 120
Db 61 SNEAGLLITSPGAILLEASPFINISSLPDININVKIEQHVLTSGKSEITLKGKDV 120
Qy 121 DOYPRLOEVSTENPLILTKLLKSIITAFAASIQESRPILTGVHIVLSNHKDFKAVAT 180
Db 121 DOYPRLOEVSTENPLILTKLLKSIITAFAASIQESRPILTGVHIVLSNHKDFKAVAT 180
Qy 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
Db 181 DSHRMSQRLITLNDTSADFDVWIPCKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
Qy 241 YTRLLEGNYPTDRLMTTEFEVFNQSLRHAMERAFILSNATQNTGVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLMTTEFEVFNQSLRHAMERAFILSNATQNTGVKLEITQNHIS 300
Qy 301 AHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFILSPVRPFTLT 360
Db 301 AHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFILSPVRPFTLT 360
Qy 361 PGDEESFIQLITPVRTN 378
Db 361 PGDEESFIQLITPVRTN 378

```

RESULT 6

```

RESULT 5
Q8K8Z9
ID Q8K8Z9 PRELIMINARY; PRT; 378 AA.
AC Q8K8Z9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Beta subunit of DNA polymerase III.
GN Name=dnaN; OrderedLocusNames=SpyM3_0002;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_taxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mamarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; AE014136; AAM78609.1; -
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
DR Complete proteome.
KW SEQUENCE 378 AA; 41861 MW; 24B19F5944A0E3B4 CRC64;

Query Match 98.3%; Score 1832; DB 2; Length 378;
Best Local Similarity 98.1%; Pred. No. 2.9e-113;
Matches 371; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MIQPSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGQISIENTIPV 60
Db 1 MIQPSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTPTGVTLTGNGQISIENTIPV 60
Qy 61 SNEAGLLITSPGAILLEASPFINISSLPDISINVKIEQHVLTSGKSEITLKGKDV 120
Db 61 SNEAGLLITSPGAILLEASPFINISSLPDININVKIEQHVLTSGKSEITLKGKDV 120
Qy 121 DOYPRLOEVSTENPLILTKLLKSIITAFAASIQESRPILTGVHIVLSNHKDFKAVAT 180
Db 121 DOYPRLOEVSTENPLILTKLLKSIITAFAASIQESRPILTGVHIVLSNHKDFKAVAT 180
Qy 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
Db 181 DSHRMSQRLITLNDTSADFDVWIPCKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
Qy 241 YTRLLEGNYPTDRLMTTEFEVFNQSLRHAMERAFILSNATQNTGVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLMTTEFEVFNQSLRHAMERAFILSNATQNTGVKLEITQNHIS 300
Qy 301 AHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFILSPVRPFTLT 360
Db 301 AHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFILSPVRPFTLT 360
Qy 361 PGDEESFIQLITPVRTN 378
Db 361 PGDEESFIQLITPVRTN 378

```

```
Q8DWN8
ID Q8DWN8 PRELIMINARY; PRT; 378 AA.
AC Q8E216;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative DNA polymerase Iii, beta subunit (SC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=SMU.02;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype c;
RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; AE014853; AAN57795.1; -
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006260; F:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc_1.
DR TIGRFAMs; TIGR00663; dnan; 1.
DR Complete proteome; Nucleotidyltransferase; Transferase.
KW Complete proteome;
SQ SEQUENCE 378 AA; 42074 MW; DP9822F73D31CAC6 CRC64;

Query Match 82.0%; Score 1527; DB 2; Length 378;
Best Local Similarity 79.9%; Pred. No. 4.7e-93;
Matches 302; Conservative 39; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MIQFSINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
Db 1 MIKFSINKVFFQLNALNATKRAISSKNAIPILSSIKIEVNSQSITLTGSGQISIENTISA 60
Qy 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEIQHVLTSGKSEITLKGKDV 120
Db 61 BEENAGLVTSGLLEAFINIVSLPDIITLDFEIEQHVLTSGKSEITLKGKDV 120
Qy 121 DQYPRQEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGVHIVLSNHNKDFKAVAT 180
Db 121 EQYPRQEVGTNNPLILTKLLKIISTAFASLQESRPILTGVHIVLTNHNKDFKAVAT 180
Qy 181 DSHRMSORLITLDNTSADLMVLPKSLRFSAVFTDDIETVEVFFSPQILFRSEHISF 240
Db 181 DSHRMSORLITLDHSSDDFDVVIPSRSLRFAAVFTDDIESVEVFFSPQILFRSEYISF 240
Qy 241 YTRLLEGNYPTDRLMTFETEVVFNQSLRHAMERAFILSNATQGTVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLGLNNFETEVVFNALRAMEAHLISNATQGTVKLEITNNQVT 300
Qy 301 AHVNSPEVGKVNEDLIDVSGSGDLTISFNPTYLIESLKAKSETVKTHFLSPVRPFTLT 360
Db 301 AHVNSPEVGKVNEDLIDSLSGNDLTISFNPTYLIEALKALKSETVTIRFISPIRPFTLT 360
Qy 361 PGDEESFIQITPVRTN 378
Db 361 PSDNSENFIQITPVRTN 378

RESULT 7
Q8E216
ID Q8E216 PRELIMINARY; PRT; 378 AA.
AC Q8E216;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE DNA polymerase Iii, beta subunit.
GN Name=dnaN; OrderedLocusNames=SAG0002;
OS Streptococcus agalactiae (serotype V);
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014191; AAM98910.1; -
DR TIGR; SAG0002; -
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; F:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
DR Complete proteome.
SQ SEQUENCE 378 AA; 42306 MW; 270D22643835C3AF CRC64;

Query Match 80.0%; Score 1491; DB 2; Length 378;
Best Local Similarity 77.5%; Pred. No. 1.1e-90;
Matches 293; Conservative 39; Mismatches 46; Indels 0; Gaps 0;

Qy 1 MIQFSINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
Db 1 MIHFSINKVFFLHALTVTKRAISHKNAIPILSTVKIEVTRDAIILTSGQISIENTIPA 60
Qy 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEIQHVLTSGKSEITLKGKDV 120
Db 61 SNEAGLLVTPGSGILLEGFFINIISLPDVTLEIEQHVLTSGKSEITLKGKDV 120
Qy 121 DQYPRQEVSTENPLILTKLLKSIIAETAFAASLQESRPILTGVHIVLSNHNKDFKAVAT 180
Db 121 DQYPRQEVGTTDPTLTLETLLKLSIINETAFAASQESRPILTGVHIVISQNKYFKAVAT 180
Qy 181 DSHRMSORLITLDNTSADLMVLPKSLRFSAVFTDDIETVEVFFSPQILFRSEHISF 240
Db 181 DSHRMSORTQLESANNFDLVFESKSLRFSAVFTDDIETVEVFFSDQMLFRSEHISF 240
Qy 241 YTRLLEGNYPTDRLMTFETEVVFNQSLRHAMERAFILSNATQGTVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLTNQFETEIFNTNLRHAMERAYLISNATQGTVRLEIQNETVS 300
Qy 301 AHVNSPEVGKVNEDLIDVSGSGDLTISFNPTYLIESLKAKSETVKTHFLSPVRPFTLT 360
Db 301 AHVNSPEVGKVNEDLTVSLKGSINISFNPTYLIESLKAVKSETVTIRFISPIRPFTLT 360
Qy 361 PGDEESFIQITPVRTN 378
Db 361 PGDETFIQITPVRTN 378
```

```

RESULT 8
Q8E723
ID Q8E723 PRELIMINARY; PRT; 378 AA.
AC Q8E723
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Beta subunit of DNA polymerase III.
GN Name=dnaN; OrderedLocusNames=gbs0002;
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Ruanio C., Buchrieser C., Chevalier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lelioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766843; CAD45647.1; -.
DR SAGAList; gbs0002; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome.
SQ SEQUENCE 378 AA; 42306 MW; 270D22643835C3AF CRC64;

Query Match 80.0%; Score 1491; DB 2; Length 378;
Best Local Similarity 77.5%; Pred. No. 1.1e-90;
Matches 293; Conservative 39; Mismatches 46; Indels 0; Gaps 0;

Qy 1 MIQPSINTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGNGQISIENTIPV 60
Db 1 MHFSINKNFELHALVTVKRAISHKNAIPILSTVKIEVTRDAIILTSGNGQISIENTIPA 60

Qy 61 SNEAGLITSPGAILLEASPFNIISSLPDISINVKIEHQHVLTSGKSEITLKGKDV 120
Db 61 SNEAGLVTNPGSILLGAGPFIINISLPDVTLEFIEHQHIVLTSGKSEITLKGKDV 120

Qy 121 DQYPRLOEVSTENPILKTLKLSIAETAFAASLQESRPILTGVHIVLSNHHKDFKAVAT 180
Db 121 DQYPRLOEWTDPTLTETLKLKLSIINETAFAASQESRPILTGVHIVISQNKYKAVAT 180

Qy 181 DSHRMSQRLITLDNTSADLMVLPKSLREFSAVFTDDIETVEVFPFSQILFRSEHISF 240
Db 181 DSHRMSQRTFQLEKANNFDLVVPSKSLREFSAVFTDDIETVEVFPFSQMLFRSENISF 240

Qy 241 YTRLLEGNYPTDRLIMTEFETEVVNTQSLRHAMERAFISNATQNTGVKLEITQNHIS 300
Db 241 YTRLLEGNYPTDRLTLTQFETEIFNTRALRHAMERAYLISNATQNTGVRLTQNETVS 300

Qy 301 AHVNSPEVGVKNEDIDIVSQSGDLTISFNPTYLSESLKAESEVTKHFLSPVPRPFLT 360
Db 301 AHVNSPEVGVKNNEEDLTVSLKGDLSNIFNPTYLSESLKAESEVTKHFLSPVPRPFLT 360

Qy 361 PGDEESFIQLTPVRTN 378
Db 361 PGDETEDFIQLTPVRTN 378

RESULT 9
DP3B_STRPN
ID DP3B_STRPN STANDARD; PRT; 378 AA.
AC O06672;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE DNA polymerase III, beta chain (SC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=SP0002;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R801;
RL Gasc A.A.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radue D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506 (2001).
CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC independent) along duplex DNA (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. POLIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000658; AAC45337.1; -.
DR EMBL; AE007318; AAK74195.1; -.
DR PIR; B95000; B95000.
DR PIR; B97872; B97872.
DR TIGR; SP0002; -.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome; DNA replication; DNA-directed DNA polymerase;
KW Transferrase.
FT CONFLICT 17 17 T -> I (in Ref. 1).
FT CONFLICT 43 43 I -> V (in Ref. 1).
FT SEQUENCE 378 AA; 42072 MW; 5749CB6E857FFE6 CRC64;
SQ SEQUENCE 378 AA; 42072 MW; 5749CB6E857FFE6 CRC64;

```

```

Query Match      74.6%; Score 1389; DB 1; Length 378;
Best Local Similarity 72.2%; Pred. No. 6, 6e-84;
Matches 273; Conservative 50; Mismatches 55; Indels 0; Gaps 0;

Cc 1 MIQPSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
Cc 1 MIHFSINKNLFQALNTTKRAISSKNAIPILSTVKIDVTNEGITLGSNGQISIENTFSQ 60
Cc 61 SNEAGLLITSPGAILLEASFFINIISLDPDISINVKIEHOHVLTSGKSEITLKGKDV 120
Cc 61 KNEADAGLLITSLGSLLEASFFINVSLSLDPDVTDFKEIQNQIVLTSGKSEITLKGKDS 120
Cc 121 DQYPRLOEVSTENPLIKTLKLSIIATAFAASLQESRPILTGTVHIVLSNHKDKFVAT 180
Cc 121 EYPRIOEISASTPLLETKLKKIINETAFATQESRPILTGTVHIVLSNHKDKFVAT 180
Cc 181 DSHRMSQRLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
Cc 181 DSHRLSQKLTLEKNSDDFDVVPISRLRSESAVFTDDIETVEIFFANNQILFRSEHISF 240
Cc 241 YTRLLEGNYPTDRLMTPEFETEVFNTQSLRHAMERAFILSNATQGTVKLEITQNHIS 300
Cc 241 YTRLLEGNYPTDRLIPTDFNTTTFNVNLRQSMERARLLSSATQGTVKLEIKDGWVS 300
Cc 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKETVKIHFSPVRPFTLT 360
Cc 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKETVKIHFSPVRPFTLT 360
Cc 361 PGDEESFIQILTPVRTN 378
Cc 361 PADTDEDFMQLITPVRTN 378

RESULT 10
DP3B_STRR6
ID DP3B_STRR6 STANDARD; PRT; 378 AA.
AC P59651;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=spr0002;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoakins J., Alborn W.E. Jr., Arnold J., Blazczak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fricz L., Fu D.-J., Fuller W., Geisinger C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
Cc -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
Cc responsible for most of the replicative synthesis in bacteria.
Cc This DNA polymerase also exhibits 3' to 5' exonuclease activity.
Cc The beta chain is required for initiation of replication once it
Cc is clamped onto DNA. It slides freely (bidirectional and ATP-
Cc independent) along duplex DNA (By similarity).
Cc -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
Cc + (DNA)(N).
Cc -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
Cc epsilon and theta chains) that associates with a tau subunit. This
Cc core dimerizes to form the POLIII' complex. POLIII' associates
Cc with the gamma complex (composed of gamma, delta, delta', psi and

```

```

chi chains) and with the beta chain to form the complete DNA
polymerase III complex (By similarity).
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (see http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
-----
Cc EMBL; AE008385; AAK98806.1; .
Cc InterPro: IPR001001; DNA_polIII_beta.
Cc Pfam; PF00712; DNA_pol3_beta_1.
Cc Pfam; PF02767; DNA_pol3_beta_2; 1.
Cc Pfam; PF02768; DNA_pol3_beta_3; 1.
Cc SMART; SM00480; POL3Bc; 1.
Cc TIGRFAMs; TIGR00663; dnan; 1.
Cc Complete proteome; DNA replication; DNA-directed DNA polymerase;
Cc Transferencease 378 AA; 42070 MW; 14CE29278184F906 CRC64;
Cc SEQUENCE 378 AA; 42070 MW; 14CE29278184F906 CRC64;
-----
Query Match      74.3%; Score 1384; DB 1; Length 378;
Best Local Similarity 72.2%; Pred. No. 1, 4e-83;
Matches 273; Conservative 49; Mismatches 56; Indels 0; Gaps 0;

Cc 1 MIQPSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
Cc 1 MIHFSINKNLFQALNTTKRAISSKNAIPILSTVKIDVTNEGVTLTSGNGQISIENTFSQ 60
Cc 61 SNEAGLLITSPGAILLEASFFINIISLDPDISINVKIEHOHVLTSGKSEITLKGKDV 120
Cc 61 KNEADAGLLITSLGSLLEASFFINVSLSLDPDVTDFKEIQNQIVLTSGKSEITLKGKDS 120
Cc 121 DQYPRLOEVSTENPLIKTLKLSIIATAFAASLQESRPILTGTVHIVLSNHKDKFVAT 180
Cc 121 EYPRIOEISASTPLLETKLKKIINETAFATQESRPILTGTVHIVLSNHKDKFVAT 180
Cc 181 DSHRMSQRLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
Cc 181 DSHRLSQKLTLEKNSDDFDVVPISRLRSESAVFTDDIETVEIFFANNQILFRSEHISF 240
Cc 241 YTRLLEGNYPTDRLMTPEFETEVFNTQSLRHAMERAFILSNATQGTVKLEITQNHIS 300
Cc 241 YTRLLEGNYPTDRLIPTDFNTTTFNVNLRQSMERARLLSSATQGTVKLEIKDGWVS 300
Cc 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKETVKIHFSPVRPFTLT 360
Cc 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKETVKIHFSPVRPFTLT 360
Cc 361 PGDEESFIQILTPVRTN 378
Cc 361 PADTDEDFMQLITPVRTN 378

RESULT 11
DP3B_LACLA
ID DP3B_LACLA STANDARD; PRT; 380 AA.
AC Q9CJL1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=LL0002;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=11403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.169701;
RX Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,

```

```

RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RL lactis ssp. lactis Ili403.";
RC Genome Res. 11:731-753(2001).
CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC independent) along duplex DNA (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. PolIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006239; AAK04100.1; -.
CC PIR; B86625; B86625.
CC InterPro; IPR001001; DNA_polIII_beta.
CC Pfam; PF00712; DNA_pol3_beta_1.
CC Pfam; PF02767; DNA_pol3_beta_2; 1.
CC Pfam; PF02768; DNA_pol3_beta_3; 1.
CC SMART; SM00480; POL3Bc; 1.
CC TIGRFAMs; TIGR00663; dnan; 1.
CC Complete proteome; DNA replication; DNA-directed DNA polymerase;
KW Transferrase.
SQ SEQUENCE 380 AA; 42264 MW; 2D54BD83932029B0 CRC64;

Query Match 56.5%; Score 1053; DB 1; Length 380;
Best Local Similarity 54.7%; Pred. No. 1.2e-61;
Matches 208; Conservative 73; Mismatches 97; Indels 2; Gaps 1;

QY 1 MIQSFINTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGTGVTGNGQISIENTIPV 60
DB 1 MIKESINKNAQNARITKQALGSKVTIPALTCLKIEVEENGITLIGSNGQISIKNPLPV 60

QY 61 SNEAGLLITSPGAILLEASFINIISLPDISINVKIEHQHVLTSGKSEITLKGKDV 120
DB 61 DNKDSMLISGTGVLLEAAFFENVVSQLPEVLTETKEQKQVLLTSGKSEITLKGKLD 120

QY 121 DOYPLQEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNHKDFKAVAT 180
DB 121 EIYPHLQISEGSSLKMKVKVLEIFETVFAVSTQENRPIFTGVHLETLSTGELKAVAT 180

QY 181 DSHRMSQRLITLDNTSADLMVLPSPKSLREFSAVTDDETVEVFPSPQILFRSEHSF 240
DB 181 LQPRNEESGFVQLITPVRTN 380

QY 181 DSHRMSQRLITLDNTSADLMVLPSPKSLREFSAVTDDETVEVFPSPQILFRSEHSF 240
DB 181 LQPRNEESGFVQLITPVRTN 380

```

```

RESULT 12
ID DP3B LACLC STANDARD; PRT; 380 AA.
AC 054376;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN Name=dnan;
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OX Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
CX NCBI_TaxID=1359;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MGI363;
RA MEDLINE=98118563; PubMed=9435243;
RA El-Karoui M., Ehrlich S.D., Gruss A.;
RT "Identification of the lactococcal exonuclease/recombinase and its
modulation by the putative Chi sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:626-631(1998).
CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC independent) along duplex DNA (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. PolIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U76424; AAC12964.1; -.
CC PIR; T30306; T30306.
CC InterPro; IPR001001; DNA_polIII_beta.
CC Pfam; PF00712; DNA_pol3_beta_1.
CC Pfam; PF02767; DNA_pol3_beta_2; 1.
CC Pfam; PF02768; DNA_pol3_beta_3; 1.
CC SMART; SM00480; POL3Bc; 1.
CC TIGRFAMs; TIGR00663; dnan; 1.
CC DNA replication; DNA-directed DNA polymerase; Transferrase.
KW SEQUENCE 380 AA; 42209 MW; 487297F3E822644 CRC64;

Query Match 56.5%; Score 1052; DB 1; Length 380;
Best Local Similarity 54.7%; Pred. No. 1.4e-61;
Matches 208; Conservative 74; Mismatches 96; Indels 2; Gaps 1;

QY 1 MIQSFINTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGTGVTGNGQISIENTIPV 60
DB 1 MIKESINKNAQNARITKQALGSKVTIPALTCLKIEVEENGITLIGSNGQISIKNPLPA 60

QY 61 SNEAGLLITSPGAILLEASFINIISLPDISINVKIEHQHVLTSGKSEITLKGKDV 120
DB 61 DNKDSMLISGTGVLLEAAFFENVVSQLPEVLTETKEQKQVLLTSGKSEITLKGKLD 120

QY 121 DOYPLQEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNHKDFKAVAT 180
DB 121 EIYPHLQISEGSSLKMKVKVLEIFETVFAVSTQENRPIFTGVHLETLSTGELKAVAT 180

QY 181 DSHRMSQRLITLDNTSADLMVLPSPKSLREFSAVTDDETVEVFPSPQILFRSEHSF 240
DB 181 LQPRNEESGFVQLITPVRTN 380

```

```

Db 181 DSHRMSQRLPLEDTLKFDPVLPKSKINSFNKFTNDEEIEIFISQSMFLQNETISY 240
Qy 241 YTRLLEGNYPTDRLLMTE--FETEVNTQSLRHAMERAFNISNATONGVTKLEITONH 298
Db 241 YSRLIEGYPNRLNLPNEADYTLDLVDAQLRHTWDRLTLTWTNGVTKLVTSGDS 300
Qy 299 ISAHVNSPEVGVKNEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRPPT 358
Db 301 VVTANSPEVGSVHEELTALSKEGNDLSISFNPEYLIDALKVIRKAPVIRFISNVRPPT 360
Qy 359 LTPGDEESFQTLITPVRTN 378
Db 361 LQPRNEESGFVLITPVRTN 380

RESULT 13
Q9L571 PRELIMINARY; PRT; 189 AA.
AC Q9L571;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA polymerase III beta subunit (EC 2.7.7.7) (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370;
RX MEDLINE=20389620; PubMed=10930754;
RA Suvorova A.N., Ferretti J.J.;
RT "Replication origin of streptococcus pyogenes, organization and
  cloning in heterologous systems.";
RL FEMS Microbiol. Lett. 189:293-297(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370;
RA Suvorov A., Ferretti J.J.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255728; AAF71536.1; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR SMART; SM00480; POL3Bc; 1.
KW Nucleotidyltransferase; Transferase.
FT NON TER 189
SQ SEQUENCE 189 AA; 20570 MW; BDDC423DBA1C441 CRC64;

Query Match 49.2%; Score 916; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 5.7e-53;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIQPSINTLFHALNTTKRAISTKNAIPILSSIKIEVTSVTLTGSGQISIENTIPV 60
Db 1 MIQPSINTLFHALNTTKRAISTKNAIPILSSIKIEVTSVTLTGSGQISIENTIPV 60
Qy 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKGSEITLKGKDV 120
Db 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKGSEITLKGKDV 120
Qy 121 DQYPRLOQVSTENPILTKLKSIIAETAFAASIQESRPILTGVHIVLSNHHKDFKAVAT 180
Db 121 DQYPRLOQVSTENPILTKLKSIIAETAFAASIQESRPILTGVHIVLSNHHKDFKAVAT 180
Qy 181 DSHRMSQRL 189
Db 181 DSHRMSQRL 189

```

```

RESULT 14
Q839Z4 PRELIMINARY; PRT; 376 AA.
ID Q839Z4
AC Q839Z4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA polymerase III, beta subunit.
GN Name=dnaN; OrderedLocusNames=EF0002;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seehadri R.,
  Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
  Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
  Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
  Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
  Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
  Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
  Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AE016947; AAC79887.1; -.
DR TIGR; EF0002; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome.
SQ SEQUENCE 376 AA; 42214 MW; 2D992530A0493B4A CRC64;

Query Match 46.4%; Score 864.5; DB 2; Length 376;
Best Local Similarity 45.9%; Pred. No. 3.8e-49;
Matches 173; Conservative 77; Mismatches 126; Indels 1; Gaps 1;

Qy 2 IQPSINTLFHALNTTKRAISTKNAIPILSSIKIEVTSVTLTGSGQISIENTIPVS 61
Db 1 MKLTVKRSVFLQELQTVQRAISSKTTIPILTVGVLSLSEDLGSLTGSNADISIESFLSKD 60
Qy 62 NENAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKGSEITLKGKDV 121
Db 61 DEKAQMTIERTGSLVLSQSRPFGEIIRKLPDPMFTMEVDNNQVAITSKADFTVNGLDAD 120
Qy 122 QYPRLOQVSTENPILTKLKSIIAETAFAASIQESRPILTGVHIVLSNHHKDFKAVAT 181
Db 121 NYPHLPVIDTQNMKLPVHLTKIISSETGFAVSMHESRPILTGVHFILENQK-LLAATD 179
Qy 182 SHRMSQRLITLDNTSADLMVLPKSLREFSAVETDDIEIVVEVFPFSQILFRSEHISFY 241
Db 180 SHRLSORVITEQVAEDFNIVIPGKSLTSLRSRSTNEEEMVEISIMENQVLFKTTWVFY 239
Qy 242 TRLEGNVPDTRLLMTTEFETEVNTQSLRHAMERAFNISNATONGVTKLEITONHISA 301
Db 240 SRLLEGNPDPNRLIPTSHNTQIEFYVPELLSALERALLSHEGNNIVRLISIPDSVVL 299
Qy 302 HVNSPEVGVKNEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRPPTLP 361
Db 300 YGNSPEIGKVEEALNENVSGEALDISFNPDYMKDALRAFGDMNITVKFLSPIRPPTLEP 359
Qy 362 GDEESFIQLITPVRTN 378

```


Db 360 LVTEGEGFIQLITPVRT 378

RESULT 16

Q74M33 PRELIMINARY; PRT; 376 AA.

AC Q74M33

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE DNA polymerase III, beta chain.

GN OrderedLocusNames=IJ0002;

OS Lactobacillus Johnsonii.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;

OC Lactobacillus.

OX NCBI_TaxID=33959;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCC 533;

RX PubMed=14966310;

RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R., Mollet B., Mercener A., Kleenhammer T., Arigoni F., Schell M.A.; "The genome sequence of the probiotic intestinal bacterium Lactobacillus Johnsonii, NCC 533.";

RT Lactobacillus Johnsonii, NCC 533.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).

DR EMBL; AE017200; AAS07982.1; -

DR InterPro; IPR001001; DNA_polIII_beta.

DR Pfam; PF00712; DNA_pol3_beta_1;

DR Pfam; PF02767; DNA_pol3_beta_2; 1.

DR Pfam; PF02768; DNA_pol3_beta_3; 1.

DR SMART; SM00480; POL3BC; 1.

DR TIGRFAMs; TIGR00663; dnan; 1.

KW Complete proteome.

SQ SEQUENCE 376 AA; 41819 MW; 33FD6403EC85AA28 CRC64;

Query Match 41.4%; Score 770.5; DB 2; Length 376;

Best Local Similarity 44.4%; Pred. No. 6.4e-43;

Matches 170; Conservative 73; Mismatches 125; Indels 15; Gaps 7;

Qy 2 IQPSINRTFIHALNTTKAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPVS 61

Db 1 MQFTINRLFLENLNNMRAISRATIPILSGIKLTLTDEMLTSGDTSIEIQIPVN 60

Qy 62 NENAGLITSPGAILLEASFFINIISSLP--DISINVKIEHQVVLTSKSEITLKGKD 119

Db 61 DD---LIVQSTGIVLPARFSEIVKLPCKDFSEVKSFQTKIV--SENTFEMGLD 115

Qy 120 VDQYPRLOEVSTENPLILTKLLKSIIAETAFAASLOESRPILTGVIHVLNKHDKAKVA 179

Db 116 ANNYPHLPETISDASFKISGKTFRFIIINETVFAVATQESRPTLTGVNFIF--NNSSIKAVA 174

Qy 180 TDSHRMSORLITLDN--TSADLMVVLPSKSLREFSAVTDDETVEVFPSPQILFRSE 236

Db 175 TDSHRUSQRIQSIENGPGQSTDL--IIPGKSLVELSRIIGSDPEITVNGENQVLFVVG 232

Qy 237 HISFYTRLLEGNYPDTRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVK--LEI 294

Db 233 NIAPYSELDDGQYPTDRLIPIESTTSVEFELPVLARSERASLLTHESNNVVKMTLDV 292

Qy 295 TQNHISAHVNSPEVGVKNVDLDIVSQSGDLTISFNPTVLIISLKAISKETVKIHLSPV 354

Db 293 QNQLVKLGDSPIGNVEEIGPKLEGDLTISFNPDYLRALRASITDSIIMNFTQPL 352

Qy 355 RPTLTPGDEESFIQLITPVRT 377

Db 353 RPFTVTPAKQDVNFTQLITPVRT 375

RESULT 17

AAS07982 PRELIMINARY; PRT; 376 AA.

AC AAS07982

Db 360 TETELDFIQLITPVRTN 376

RESULT 15

Q890K7 PRELIMINARY; PRT; 379 AA.

AC Q890K7

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE DNA-directed DNA polymerase III, beta chain (EC 2.7.7.7).

GN Name=dnan; OrderedLocusNames=lp_0002;

OS Lactobacillus plantarum.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;

OC Lactobacillus.

OX NCBI_TaxID=1590;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCIMB 8826 / WCFS1;

RX MEDLINE=22480296; PubMed=12565666;

RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D., Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M., Fliers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A., Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B., De Vos W.M., Siezen R.J.; "Complete genome sequence of Lactobacillus plantarum WCFS1.";

RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).

DR EMBL; AL935252; CAD62704.1; -

DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.

DR GO; GO:0016740; F:transfrase activity; IEA.

DR GO; GO:0006260; P:DNA replication; IEA.

DR InterPro; IPR001001; DNA_polIII_beta.

DR Pfam; PF00712; DNA_pol3_beta_1;

DR Pfam; PF02767; DNA_pol3_beta_2; 1.

DR Pfam; PF02768; DNA_pol3_beta_3; 1.

DR SMART; SM00480; POL3BC; 1.

DR TIGRFAMs; TIGR00663; dnan; 1.

KW Complete proteome; DNA-directed DNA polymerase;

KW Nucleotidyltransferase; Transferase.

SQ SEQUENCE 379 AA; 41520 MW; 13EA997B81E73717 CRC64;

Query Match 44.6%; Score 830; DB 2; Length 379;

Best Local Similarity 45.9%; Pred. No. 7.4e-47;

Matches 174; Conservative 73; Mismatches 128; Indels 4; Gaps 3;

Qy 2 IQPSINRTFIHALNTTKAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPVS 61

Db 1 MKFTINRSAPFIELNPNVQSAISKTPIPLTGKLDVNTDAITLGTSGDADISIIETIPAS 60

Qy 62 NENAGLITSPGAILLEASFFINIISSLPDISINVKIEHQVVLTSKSEITLKGKVD 121

Db 61 DDNTLVVEDAGSIVLPARFSEIVKLPEDTNTVNVVDFQITQSGAASFTINGLDPE 120

Qy 122 QYPRLOEVSTENPLILTKLLKSIIAETAFAASLOESRPILTGVIHVLNKHDKAKVATD 181

Db 121 NYPLHPEIDTNTITLADGVKELIQTVIAVNSQESRPILTGVIHVLN--GEPLAVATD 179

Qy 182 SHRMSORLITL--DNTSADLMVVLPSKSLREFSAVTDDETVEVFPSPQILFRSEHISF 240

Db 180 SHRLSORRIKLPEANNANTDVITPGKSLTELSRMIGDNNPDVQMRLENQVLFVLGNTSF 239

Qy 241 YTRLLEGNYPDTRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLEI--TQNH 298

Db 240 YSRLLGNYPDTRSLIPKESNTTVLSAPALSAERASLLSHESNNVFRVSNPTDKT 299

Qy 299 ISAHVNSPEVGVKNVDLDIVSQSGDLTISFNPTVLIISLKAISKETVKIHLSPVRPFT 358

Db 300 ITIFGNSPDVGEVTEQLPTDLSGDELEISFNPDYMKELRSFGQAMIKISFTMALRPFT 359

Qy 359 LTPGDEESFIQLITPVRT 377

DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE DNA polymerase III, beta chain.
GN L30002.
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Fridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Fittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollat B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.,
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533".
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517 (2004).
DR EMBL: A0017200; AAS07982.1; -.
SQ SEQUENCE 376 AA; 41819 MW; 33FD6403EC85AA28 CRC64;
Query Match 41.4%; Score 770.5; DB 2; Length 376;
Best Local Similarity 44.4%; Pred. No. 6.4e-43;
Matches 170; Conservative 73; Mismatches 125; Indels 15; Gaps 7;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGNQISIENTIPVS 61
DB 1 MQFTINRNLFLNLANMAIRAISRATIPILSGIKLTLTDEMLTGTSDTISIEIPIVN 60
QY 62 NENAGLLITSPGAILLEAFINIISSLP--DISINVKIEHQVLTSGSKSEITLKGKD 119
DB 61 DD---LIVOSTGIVLPARFSEIVKVLPGKDFSEVKESPTKIV--SENTEPMINGLD 115
QY 120 VDQYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDFKAVA 179
DB 116 ANNYPHLPEISTDASFKISGKTFREINETVAVATQESRPTLTGVNFI--NNSIKAVA 174
QY 180 TDSHRMSQRLITLDN---TSADLMVLPKSLRFSVAFTDDIETVVFSPQILFRSE 236
DB 175 TDSHRLSQRLISLENGPQTSTDL--IIPKSLVLSRIIGSDPEITVNPGENQVLPVVG 232
QY 237 HISFYTRLLEGNYPTDRLMTETETEVVENTQSLRHAMERAFISNATONGTVK--LEI 294
DB 233 NIAFSRLDGGYPTDRLITESTTVEFELPVLARSLERASLLTHESRNNVVKMTLDV 292
QY 295 TONHISAHVNSPEVGKVNEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFILSPV 354
DB 293 QNQLVKLQDSPSEIGNVEEIGFKNLEGDGLTISFNPDYLRALRASITDSDIINFTQPL 352
QY 355 RPTLTGDEEESFQILITPVRT 377
DB 353 RPTVTPAKQDVNFTQILITPVRT 375
RESULT 18
DP3B_STAEF
ID DP3B STAEF STANDARD; PRT; 377 AA.
AC Q8CQK6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN Name=dnaN, OrderedLocNames=SE00002;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12850922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,

RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593 (2003).
CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC independent) along duplex DNA (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. POLIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: A0016744; AAO03599.1; -.
DR InterPro: IPR001001; DNA_polIII_beta.
DR Pfam: PF00712; DNA_pol3_beta; 1.
DR Pfam: PF02767; DNA_pol3_beta_2; 1.
DR Pfam: PF02768; DNA_pol3_beta_3; 1.
DR TIGRFAMs: TIGR00663; dnaN; 1.
KW Complete proteome; DNA replication; DNA-directed DNA polymerase;
KW Transference.
SQ SEQUENCE 377 AA; 42101 MW; 4189F2AFDF3BCBDD CRC64;
Query Match 40.2%; Score 748.5; DB 1; Length 377;
Best Local Similarity 40.5%; Pred. No. 1.9e-41;
Matches 153; Conservative 86; Mismatches 136; Indels 3; Gaps 3;
QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGNQISIENTIPV 60
DB 1 MBEFTIKRDYFINQLANDTLKAIISPTTLPITGKIDAKENEVILTGSDESISEITIPK 60
QY 61 SNENAGLL-ITSPGAILLEAFINIISSLPDISINVKIEHQVLTSGSKSEITLKGKD 119
DB 61 QVDGEEIVEITGTVVLPGRFFVDIIKLPGKVEKLTSTNEQFTLTITSGHSEFNLSGLD 120
QY 120 VDQYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDFKAVA 179
DB 121 PDQYPLLPEVSRDDAIQLSVKVLKNIIAQTNFAVSTSETPVLTGVNWLQDN-ELICTA 179
QY 180 TDSHRMSQRLITLDNTSADLMVLPKSLRFSVAFTDDIETVVFSPQILFRSEHIS 239
DB 180 TDSHRLAVRLKLEDESENKNVPIPGKALSELNKMISDSDDEDIDIFFASQVLFVGNIN 239
QY 240 FYTRLLEGNYPTDRLIMTEFEFEVFNQSLRHEMERAFISNATONGTVKLEITQNH 299
DB 240 FISRLLEGHPDTRFLFPENIEIKGINNGDFHAI DRASLARLEGNNVILKSTGNELV 299
QY 300 SAHNSPEVGKVNEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFILSPVFPF 359
DB 300 ELSSTSPFETGVKEVNDVEGNNLKISFNKYNMALKAINDEVEVEFFCTMKPFIL 359
QY 360 TPGDEEESFQILITPVRT 377
DB 360 KPKD-DDSVTQLILPRT 376
RESULT 19
DP3B_STAAM

ID DP3B STAA: STANDARD; PRT; 377 AA.
 AC P50029;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE DNA polymerase III, beta chain (EC 2.7.7.7).
 GN Name=dnaN; OrderedLocNames=SAV0002, SA0002, MW0002;
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699);
 OS Staphylococcus aureus (strain N315);
 OS Staphylococcus aureus (strain MW2);
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=156878, 158879, 196620, 1280;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
 RA Kanchisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MW2;
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA.";
 RL Lancet 359:1819-1827(2002).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=YB886;
 RX MEDLINE=95206242; PubMed=7898435;
 RA Alonso J.C., Fisher L.M.;
 RT "Nucleotide sequence of the recF gene cluster from Staphylococcus
 RT aureus and complementation analysis in Bacillus subtilis recF
 RT mutants.";
 RL Mol. Gen. Genet. 246:680-686(1995).
 CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
 CC responsible for most of the replicative synthesis in bacteria.
 CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
 CC The beta chain is required for initiation of replication once it
 CC is clamped onto DNA, it slides freely (bidirectional and ATP-
 CC independent) along duplex DNA (By similarity).
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
 CC epsilon and theta chains) that associates with a tau subunit. This
 CC core dimerizes to form the PolIII' complex. PolIII' associates
 CC with the gamma complex (composed of gamma, delta, delta', psi and
 CC chi chains) and with the beta chain to form the complete DNA
 CC polymerase III complex (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP003358; BAB56164.1; -;
 CC EMBL; AP003129; BAB41218.1; -;
 CC EMBL; AP004822; BAB93867.1; -;

DR PIR, S54708; S54708
 DR SWISS-2DPAGE; P99103; STAA.
 DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta_1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3BC; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 KW Complete proteome; DNA replication; DNA-directed DNA polymerase;
 KW Transference.
 SQ SEQUENCE 377 AA; 41913 MW; 0A985EF94E044FBC CRC64;
 Query Match 39.2%; Score 730.5; DB 1; Length 377;
 Best Local Similarity 39.3%; Pred. No. 2.9e-40;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIOFSINRTLFHALNTTKRAISNAIPILSSIKIEVSTGVTLTGNGOISIENTIPV 60
 DB 1 MMEFTIKRDYFITQLNDTLKASIPRTTLPILTGKIDAKEHEVILTGSDSSEISITIPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISLSPDISINVKIEIQHQVLTSGKSEITLKGD 119
 DB 61 TVDGEDIVNISETSGSVLPGRFFVDIIKKLPKDKVLTSTNEQFQTLTITSGHSEFNLSGLD 120
 QY 120 VDOYPRLOEVSTENPLILTKLSITAEATAFAASLOESRPILTGCVHIVLSNHHKDFKAVA 179
 DB 121 PDQYPLLPQVSRDAIQLSVKVKNVIAQTNFVSTSETRPVLTGVNVLIOEN-ELICTA 179
 QY 180 TDSHRMSORLITLNDTSADLMVPLPSKSLRFSFAVFTDDIETVEVFFSPSQILFRSEHIS 239
 DB 180 TDSHRLAVRKLQLEDVSENKNVLPGRKALAEINKMSDNEEDIDIFFASQVLFKVGNNV 239
 QY 240 FYTLLEGNYPTDRLMTETETEVFNFTQSLRHAMERAFILSNATQNGTVKLEITQNH 299
 DB 240 FISRLLEGHYPTDTRLFPENYEIKLSDNGEFYHAIDRASLLAREGGNNVILKSTGDDVV 299
 QY 300 SAHVSNEPVGKNWEDLDIVSQSGDLTISNPTVLTSLKAIKSETVKIHLSPVRPRTL 359
 DB 300 ELSSTSPETGTVKREVDANDVEGSLKISFNKYMMDALKADINDEVEFEFGTMKPFIL 359
 QY 360 TP-GDESESFQLITPVRT 377
 DB 360 KPKGDD-DSVTQLILPRT 376
 RESULT 20
 Q6GD88 PRELIMINARY; PRT; 377 AA.
 AC Q6GD88;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE DNA polymerase III, beta chain (EC 2.7.7.7).
 GN ORFNames=SA0002;
 OS Staphylococcus aureus subsp. aureus MSSA476.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282459;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MSSA476;
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 RA James K.D., Kennard N., Line A., Mayes R., Moule S., Mungall K.,
 RA Ormond D., Quail M.A., Rabbintowich E., Rutherford K., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:
 RT evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571857; CAG41774.1; -;

DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta; 1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3BC; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 KW Nucleotidyltransferase; Transferase.
 SQ SEQUENCE 377 AA; 41913 MW; OA985EF94E044FBC CRC64;
 Query Match 39.2%; Score 730.5; DB 2; Length 377;
 Best Local Similarity 39.3%; Pred. No. 2.9e-40;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIQSFINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGOISIENTIPV 60
 DB 1 MMEFTIKRDYFITQLNDTLKALSPRTTLPILTKIDAKEHEVILTSGSDSISIEITIPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPDISINVKIEIQHVLTSGSKSEITLKGD 119
 DB 61 TVDGEDIVNSETGSVLPGRFFVDIIKKLPKGVKLTSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDOYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVHVLNSHKDFKAVA 179
 DB 121 PDQYPLLPQVSRDDAIQLSVKLVKNVIAQTAVSTSETRPVLTVGNWLIQEN-ELICTA 179
 QY 180 TDSHRMSQRLITLNTSADLMVLPKSLRFRFSVFTDDIETVEVFFSPQILFRSEHIS 239
 DB 180 TDSHRLAVRKLQLEDVSENKNVILPGKALAEINKIMSDNEEDIDIFFASNQVLPKGVN 239
 QY 240 FYTRLLEGNYPTDRLLMTETEVVNTQSLRHAMERAFILSNATONGTVKLEITONHI 299
 DB 240 FISRLLEGHPDTRLPENYEIKLSIDNGEFYHAIDRASLLAREGGNNVILKSTGDDV 299
 QY 300 SAHNSPEVGVKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRPPTL 359
 DB 300 ELSSTSPSIGVTKEVDANDVEGSLKISFNKYMMDALKDAINDNEVEVEFFGTMKPPIL 359
 QY 360 TP-GDEESFIQLITPVRT 377
 DB 360 KPKGDD--DSVTQLILPIRT 376
 RESULT 21
 Q6GKU3 PRELIMINARY; PRT; 377 AA.
 ID AC Q6GKU3
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DE DNA polymerase III, beta chain (EC 2.7.7.7).
 GN Name=dnaN; ORFNames=SAR0002;
 OS Staphylococcus aureus subsp. aureus MRSA252.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MRSA252;
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corron C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
 RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:
 RT evidence for the rapid evolution of virulence and drug resistance."
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571856; CAG39030.1; -.
 DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta; 1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.

DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3BC; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 KW Nucleotidyltransferase; Transferase.
 SQ SEQUENCE 377 AA; 41913 MW; OA985EF94E044FBC CRC64;
 Query Match 39.2%; Score 730.5; DB 2; Length 377;
 Best Local Similarity 39.3%; Pred. No. 2.9e-40;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIQSFINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGOISIENTIPV 60
 DB 1 MMEFTIKRDYFITQLNDTLKALSPRTTLPILTKIDAKEHEVILTSGSDSISIEITIPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPDISINVKIEIQHVLTSGSKSEITLKGD 119
 DB 61 TVDGEDIVNSETGSVLPGRFFVDIIKKLPKGVKLTSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDOYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVHVLNSHKDFKAVA 179
 DB 121 PDQYPLLPQVSRDDAIQLSVKLVKNVIAQTAVSTSETRPVLTVGNWLIQEN-ELICTA 179
 QY 180 TDSHRMSQRLITLNTSADLMVLPKSLRFRFSVFTDDIETVEVFFSPQILFRSEHIS 239
 DB 180 TDSHRLAVRKLQLEDVSENKNVILPGKALAEINKIMSDNEEDIDIFFASNQVLPKGVN 239
 QY 240 FYTRLLEGNYPTDRLLMTETEVVNTQSLRHAMERAFILSNATONGTVKLEITONHI 299
 DB 240 FISRLLEGHPDTRLPENYEIKLSIDNGEFYHAIDRASLLAREGGNNVILKSTGDDV 299
 QY 300 SAHNSPEVGVKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRPPTL 359
 DB 300 ELSSTSPSIGVTKEVDANDVEGSLKISFNKYMMDALKDAINDNEVEVEFFGTMKPPIL 359
 QY 360 TP-GDEESFIQLITPVRT 377
 DB 360 KPKGDD--DSVTQLILPIRT 376
 RESULT 22
 Q8EU87 PRELIMINARY; PRT; 378 AA.
 ID AC Q8EU87
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE DNA-directed DNA polymerase III beta chain (EC 2.7.7.7).
 GN Name=dnaN; OrderedLocusNames=OB00002;
 OS Oceanobacillus iheyensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831;
 RX MEDLINE=2220767; PubMed=12235376;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; AP004593; BAC11958.1; -.
 DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006260; F:DNA replication; IEA.
 DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta; 1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3BC; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 KW Complete proteome; DNA-directed DNA polymerase;

```

KW NucleotideIdyltransferase; Transferase.
SQ SEQUENCE 378 AA; 42231 MW; BABACD30AE3BD962 CRC64;

Query Match 38.8%; Score 722.5; DB 2; Length 378;
Best Local Similarity 41.1%; Pred. No. 9.8e-40;
Matches 157; Conservative 85; Mismatches 129; Indels 11; Gaps 7;

QY 2 IQFSINRILFHALNTTKRAISTKNAIPILSSIKIEVTSSTGTLTSGNGQISIENTIPVS 61
Db 1 MRFTIQDKLINGVNMKAISARTVPIILTKMKIEVKNHGVTLTGSDSDISIEYIPI- 59
QY 62 NENAGLLIT---SPCAILLEAFNIISSLPDISINVKIEIOHVVLTSGKSEITLKGK 118
Db 60 -EEDGIVHENIEEETIILQAKYFDPVVKLPSTVDIVDDQLNVRTSGRAEENLNGQ 118
QY 119 DVDQYPRLOQVSTENPLILKTLKLSIIAETAFASLOESRPILTVGHVILSNHDKRAV 178
Db 119 SAEEYQPKVQWSTENFELPIDLLAKSMIKQVFAVSTWETRPILTVGNLKLVDN-SLSFT 177
QY 179 ATDSHRMSORLITLNTSADL-MVVLPSKSLREFSAVFTDDIETVEVFFSPSOLFRSEH 237
Db 178 ATDSHRLARREIPVSNAPIEISQIVVPGKSLMELNKLIGDSEETVEISVTNNQILFRTH 237
QY 238 ISFYTRLLEGNYPDTRLMTTEFETEVFNQSLRHAMERAFILSNATQNGTVKLEITQN 297
Db 238 LNFLSLRLOGNPETSRILPEOSKTKIQLKTKELGTIDRASLLAKEERNNVKFNAPGN 297
QY 298 H-ISAHVNSPEVGVKNEDLDIVSQSGDLTTSFNPTYLIESKAISKSETVKIHFLSPVRP 356
Db 298 SMIEISSNSPEVGVNVEITADQMGEDVKISFSKYMIDALKALEYDEVQIEFTGMRP 357
QY 357 FTLTP-GDEESFIQLITPVRT 377
Db 358 FIIRPVGD--DSILQILPVRT 377

RESULT 23
Q92FV1 PRELIMINARY; PRT; 381 AA.
AC Q92FV1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA polymerase III, beta chain.
GN Name=dnaN; OrderedLocusNames=lin0002;
OC Bacteria; Firmicutes;
OC NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madseno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95235.1; -.
DR PIR; AC1433; AC1433.
DR ListList; LIN0002; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001001; DNA polIII_beta.

DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta_2; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3BC; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome.
SQ SEQUENCE 381 AA; 42497 MW; 977614257898F46D CRC64;

Query Match 38.1%; Score 710; DB 2; Length 381;
Best Local Similarity 39.4%; Pred. No. 6.7e-39;
Matches 150; Conservative 92; Mismatches 133; Indels 6; Gaps 5;

QY 2 IQFSINRILFHALNTTKRAISTKNAIPILSSIKIEVTSSTGTLTSGNGQISIENTIP-V 60
Db 1 MKFVIERDLRQVAVNEVTRAIASRTTIPILTGKIVNVDEGVTLTGSDSDISIEAFIPLI 60
QY 61 SNEAGLLITSPGAILLEASFFNIISSLPDISINVKIEIOHVVLTSGKSEITLKGKDV 120
Db 61 ENDEVIVEVESFGGIVLQSKYFGDIVRRLPENVEIEVTNYQTNISSGQASFTLGLDP 120
QY 121 DQYPRLOQVSTENPLILKTLKLSIIAETAFASLOESRPILTVGHVILSNHDKRAV 180
Db 121 MEYKPLPEVTDGKNIKIPINVKNIIQRTVFAVSAIEVRPVLTVGNWIIKDNK-LSAVAT 179
QY 181 DSHRMSORLITLNTSADL-MVVLPSKSLREFSAVFTDDIETVEVFFSPSOLFRSEHIS 239
Db 180 DSHRLARREIPLETNIDEEYNIVIPGKSLAELNKILDDASESIEMTLANNQILFKKOLL 239
QY 240 FYTRLLEGNYPDTRLMTTEFETEVFNQSLRHAMERAFILSNATQNGTVKLEITQN-H 298
Db 240 FYSRLLEGSPDTSRLIPTDTKSELVINSRAFLQIDAIDRASLLARENRNVNKLMTLENGQ 299
QY 299 ISAHVNSPEVGVKNEDLDIVSQSGDLTTSFNPTYLIESKAISKSETVKIHFLSPVRPFT 358
Db 300 VEVSNSPEVGVNSENVPFSQSTGEEIKISFNGKYMMDALRAFEAGDDIQISFGTMRPFEV 359
QY 359 LTPGD--EESFIQLITPVRT 377
Db 360 LRPKDATNPNEILQILPVRT 380

RESULT 24
Q8YAW1 PRELIMINARY; PRT; 381 AA.
AC Q8YAW1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA polymerase III, beta chain.
GN Name=dnaN; OrderedLocusNames=lmo0002;
OC Bacteria; Firmicutes; Bacillales;
OC NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madseno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591973; CAC98217.1; -.
DR PIR; AC1432; AC1432.
DR ListList; LMO0002; -.

```

GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
GO; GO:0003677; F:DNA binding; IEA.
GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
GO; GO:0006260; P:DNA replication; IEA.
InterPro; IPR001001; DNA_polIII_beta.
Pfam; PF00712; DNA_pol3_beta_1.
Pfam; PF02767; DNA_pol3_beta_2; 1.
Pfam; PF02768; DNA_pol3_beta_3; 1.
SMART; SM00480; POL3BC; 1.
TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome; Nucleotide transferase; Transferase.
SQ SEQUENCE 381 AA; 42429 MW; 49F34BBEDBF7844 CRC64;

Query Match 37.7%; Score 703; DB 2; Length 381;
Best Local Similarity 38.8%; Pred. No. 1.9e-38;
Matches 148; Conservative 93; Mismatches 134; Indels 6; Gaps 5;

QY	2	IQFSINRTLFHIALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGOISIENTIP-V	60
DB	1	MKFVIERDLVQAQNEVTRAIARTIPLTGKIVNDEGVLITGSDSISIAFIPI	60
QY	61	SNNAGLLITSPGAILLEASFFINISSLIPDISINVKEIQHVLTSGKSEITLKGKV	120
DB	61	ENDEVIVEVESFGIVLOSKYFGDIVRRLPENVEIVTSNYQTNISSQASFTINGLDP	120
QY	121	DQVRLQEYSTENPLITKLKLSIIAETAPAAQSQRPIITGVHIVLNHKDFKAVAT	180
DB	121	MEYPKLPEVDGTGIKIPINLVKNIVROTFAVAIEVRPVLTGVNWIIKENK-LSAVAT	179
QY	181	DSHRMSQRILTLD-NTSADLMVLPKSRLRFSFAVTTDDIETVEVFPSQILFRSEHS	239
DB	180	DSHRLALREIPLETDIDEYNIVIPKSLSELNKLLDASESIEMTLANNQILFKDKLL	239
QY	240	FYTRLLEGNYPTDRILLMTEPETEVFNTOQLRHAMERAFILSNATONGTVKLEITON-H	298
DB	240	FYSRLLEGSPYDTSLIPTDKSELVINSKFLAQDAIRASLARERNNVVKMLTENGO	299
QY	299	ISAHNVPKVGKVEDLIDVSGSDLTISFNPTYILIESLKAKSETVKIHFLSPVRPFT	358
DB	300	VEVSSNSPEVGNVSNFVSQSFTGEIKISFNGKYMDALRAFEGDDIQISFSGTMPPFV	359
QY	359	LTPGD--BEESFIQITPVRT 377	
DB	360	LRPKDAANPNELIQLITPVRT 380	

RESULT 25
Q725G9 PRELIMINARY; PRT; 381 AA.

ID	Q725G9	PRELIMINARY; PRT; 381 AA.
AC	AAT02792	
DT	10-MAY-2004	(TREMBLrel. 27, Created)
DT	10-MAY-2004	(TREMBLrel. 27, Last sequence update)
DE	10-MAY-2004	(TREMBLrel. 27, Last annotation update)
DN	DNA polymerase III, beta subunit (EC 2.7.7.7).	
GN	dnaN OR LMOF2365.0002	
OS	Listeria monocytogenes str. 4b F2365.	
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria;	
OX	Listeria monocytogenes.	
NCBI_TaxID=265669;		
[1]		
RN	SEQUENCE FROM N.A.	
RP	STRAIN=4b F2365;	
RC	PUBMed=15115801;	
RX	Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,	
RA	Kolonay J.F., Raeko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,	
RA	Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,	
RA	Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,	
RA	Haft D.H., Selengut J., Van Aken S.E., Khouri H.W., Fedorova N.,	
RA	Forberger H.A., Tran B., Kathariou S., Wonderling L.D., Uhlich G.A.,	
RA	Bayles D.O., Luchansky J.B., Fraser C.M.,	
RT	"Whole genome comparisons of serotype 4b and 1/2a strains of the food-	
RT	borne pathogen Listeria monocytogenes reveal new insights into the	
RT	core genome components of this species."	
RL	Nucleic Acids Res. 32:2386-2395(2004).	
EMBL; AE017322; AAT02792.1; -		

KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 381 AA; 42429 MW; 49F34BBDBFB7844 CRC64;

Query Match 37.7%; Score 703; DB 2; Length 381;
Best Local Similarity 38.8%; Pred. No. 1.9e-38;
Matches 148; Conservative 93; Mismatches 134; Indels 6; Gaps 5;

QY 2 IQFSINRFLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 60
DB 1 MKFVIERDLVQAVNEVTRAISSARTTIPILTGKIVNDEGVTLTGSDSDISIEAFIPLI 60
QY 61 SNEAGLLITSPGAILLEASFFINISSLDPDISINVKIEIQHVLTSGKSEITLTKGKV 120
DB 61 ENDEVIVEVSGGIVLQSKVFGDVRRLPENVEIETSTQNTWISSGQASFTLNGUDP 120
QY 121 DQYPRLOEVSTENPLILTKLKSIIAETAFAASLOESRPILTGTVHIVLSNHHKDFKAVAT 180
DB 121 MEYPKLPVTDGKTIKIPINVLKNIVRTQVFAVSAIEVRPVLTVGNWIIKENK-LSAVAT 179
QY 181 DSHRMSQRLITLD-NTSADLM--VVLPSKSLRBFSAVFTDDIETVEVFPSPQILFRSEHIS 239
DB 180 DSHRLALREIPLTDEIYINIVIPGKSLNELSKILDSEEMVDIIVITEVQVLFRTKH 238
QY 240 FYTRLLEGNYPTDRLMTETEVNTQSLRHAMERAFILSNATONGTVKLEITON-H 298
DB 240 FYSLREGSYPTSLITDTKSELVINSKFLQADIRASLLARENKNVIMKMTLENGQ 299
QY 299 ISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRPFT 358
DB 300 VEVSSNSPEVGNSVNFVSQSFTGEEIKISFNGKYMMDALRAFEAGDDIQTISFGTMRFPV 359
QY 359 LTPGD--BEESFIQILTPVRT 377
DB 360 LRPKDAANPNEILQILTPVRT 380

RESULT 27
QY Q6HQ02 PRELIMINARY; PRT; 379 AA.
AC Q6HQ02;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit (EC 2.7.7.7).
GN Name=dnaN; ORFNames=BT9727_0002;
OS Bacillus thuringiensis serovar konkukian str. 97-27.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus thuringiensis serovar konkukian.
OX NCBI_TaxID=281309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.,
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT58887.1;
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta; 2; 1.
DR Pfam; PF02768; DNA_pol3_beta; 3; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 379 AA; 42590 MW; 9A26C48CA8029323 CRC64;

Query Match 37.3%; Score 695; DB 2; Length 379;
Best Local Similarity 38.8%; Pred. No. 6.6e-38;
Matches 148; Conservative 85; Mismatches 140; Indels 8; Gaps 7;

QY 2 IQFSINRFLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 61
DB 1 MRFSIQKDYLVRSQVDVKNKAVSSRTTIPILTGKIVVATEEGVTLTGSDADISIESFIPVE 60

KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 379 AA; 42529 MW; A413B2CFD197C439 CRC64;

Query Match 37.2%; Score 693; DB 2; Length 379;
Best Local Similarity 38.6%; Pred. No. 8.9e-38;
Matches 147; Conservative 86; Mismatches 140; Indels 8; Gaps 7;

QY 2 IQFSINRFLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 61
DB 1 MRFTIQKDYLVRSQVDVKNKAVSSRTTIPILTGKIVVATEEGVTLTGSDADISIESFIPVE 60
QY 62 NENAGLL-ITSPGAILLEASFFINISSLDPDISINVKIEIQHVLTSGKSEITLTKGKV 119
DB 61 ENGKEIVEVKQSGSIVLQAKYFSEIVKLPKETVEI-SVENHLMTKITSGSEFNGLND 119
QY 120 VDQYPRLOEVSTENPLILTKLKSIIAETAFAASLOESRPILTGTVHIVLSNHHKDFKAVA 179
DB 120 SAEYPLLPOLIEEHVFKIPTDOLLKHMIRQTVFAVSTSETRPILTGVMNKNVY-SELTGIA 178
QY 180 TDSHRMSQRLITLDNTSADLM--VVLPSKSLRBFSAVFTDDIETVEVFPSPQILFRSEH 237
DB 179 TDSHRLALRAKIEGIVDEFOANVVIPOKSLNELSKILDSEEMVDIIVITEVQVLFRTKH 238
QY 238 ISFTYRLLEGNYPTDRLMTETEVNTQSLRHAMERAFILSNATONGTVKLE-ITQ 296
DB 239 LLFFSRLEGNYPTDRLTIPAESKTDIPVNTKEFLQADIRASLLARDGNVNVKLTSTLEQ 298
QY 297 NHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHLSPVRP 356
DB 299 AMLSISNSPBGKVVVEVQCEKVDGELKISFSAKYMMDALKALDSTEIKISFTGAMRP 358
QY 357 FTLTPGDEEESFIQILTPVRT 377
DB 359 FLIRTVN-DESIQILTPVRT 378

RESULT 28
QY Q73FK4 PRELIMINARY; PRT; 379 AA.
AC Q73FK4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit (EC 2.7.7.7).
GN Name=dnaN; OrderedLocusNames=BCE0002;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RL adaptations and a large plasmid related to Bacillus anthracis pXO1.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017264; AAS38938.1;
DR TIGR; BCE0002;
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta; 2; 1.
DR Pfam; PF02768; DNA_pol3_beta; 3; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 379 AA; 42529 MW; A413B2CFD197C439 CRC64;

Db 120 SAEPYLLPQIEEHVFKIPTDLLKHMIRQTVFAVSTSETPILTVGNWKYN-SELTGTA 178
QY 180 TDSHRMSORLITLNTSADLM--VVLPSKSLRFSFAVFTDDIEVTFVFFSPQILFRSEH 237
Db 179 TDSHRLALRKAKIEGIADEFOANVVIPEGKSLNELSKILDESEEMVDIVITEYQVLFRTKH 238
QY 238 ISFYTRLLEGNYPTDRLMLTEPETEVFNQTSRLHAMEAFILSNATONGTVKLE-ITQ 296
Db 239 LFFSRLLEGNYPTDRLIPAESKTDIFVNTKEFLQAIADRASLLARDGRNVVKLSTLEQ 298
QY 297 NHISAHVNSPEVKGKVEDLIVSQSGDLTISFNPTLYIESLKAISKSETVKIHFLSPVRP 356
Db 299 AMLEISSNSPEIGKVVEEQCEKVDGEELKISFSKAKYMDALKALDSTEIKISFTGMARP 358
QY 357 FTLTGPGDERESFQILTPVRT 377
Db 359 FLIRTVN-DESIQILTPVRT 378

RESULT 29
AAS38938
ID AAS38938 PRELIMINARY; PRT; 379 AA.
AC AAS38938
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit (EC 2.7.7.7).
GN DNAN OR BC00002.
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L., Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F., Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXOI.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017264; AAS38938.1; -;
DR TIGR; BC00002; -;
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 379 AA; 42529 MW; A413B2CFD197C439 CRC64;

Query Match 37.2%; Score 693; DB 2; Length 379;
Best Local Similarity 38.6%; Pred. No. 8, 9e-38;
Matches 147; Conservative 86; Mismatches 140; Indels 8; Gaps 7;
QY 2 IQFSINRILFIHALNTTKRAISTKNAIPILSIKIEVTSTGTTLTGSNGQISIENTIPVS 61
Db 1 MRFTIQDYLVRSDQVDMKAVSSRTTIPILTGKIVKATEEGVTLTGSADISIESFIPVE 60
QY 62 NENAGLL-ITSPGATLLPESAFINISLIPDISINVKIEHOV-ITSGKSEITLKGKD 119
Db 61 ENGEIVSVKSGSGLVQAKTFSFELVKKPKETVEI-SVENHMLTKITSGKSEFNGLD 119
QY 120 VDQYPRLOEVSTENPLILKTKLKSIIAETAFAASLQESRPILTGVIHLSNHHKDFKAVA 179
Db 120 SAEPYLLPQIEEHVFKIPTDLLKHMIRQTVFAVSTSETPILTVGNWKYN-SELTGTA 178
QY 180 TDSHRMSORLITLNTSADLM--VVLPSKSLRFSFAVFTDDIEVTFVFFSPQILFRSEH 237
Db 179 TDSHRLALRKAKIEGIADEFOANVVIPEGKSLNELSKILDESEEMVDIVITEYQVLFRTKH 238
QY 238 ISFYTRLLEGNYPTDRLMLTEPETEVFNQTSRLHAMEAFILSNATONGTVKLE-ITQ 296
Db 239 LFFSRLLEGNYPTDRLIPAESKTDIFVNTKEFLQAIADRASLLARDGRNVVKLSTLEQ 298
QY 297 NHISAHVNSPEVKGKVEDLIVSQSGDLTISFNPTLYIESLKAISKSETVKIHFLSPVRP 356
Db 299 AMLEISSNSPEIGKVVEEQCEKVDGEELKISFSKAKYMDALKALDSTEIKISFTGMARP 358

QY 357 FTLTGPGDERESFQILTPVRT 377
Db 359 FLIRTVN-DESIQILTPVRT 378
RESULT 30
Q81W34
ID Q81W34 PRELIMINARY; PRT; 379 AA.
AC Q81W34; Q61537; Q6KYT0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2004 (TReMBLrel. 26, Last annotation update)
DE DNA polymerase III, beta subunit.
GN Name=dnan-1; OrderedLocusNames=BA0002, BAS0002; ORFNames=GBAA0002;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=1271629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T., Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R., Holtzapfel E.K., Oekstad O.A., Helgason E., Ristone J., Wu M., Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L., DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H., Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D., Berty K.J., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F., Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O., Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RA "The genome sequence of Bacillus anthracis and comparison to closely related bacteria.";
RT Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B., Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L., Fraser C.M.;
RT "Bacillus anthracis comparative genomics";
RN Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Stearns;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K., Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R., Richardson P., Rubin E., Tice H.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017024; AAP24060.1; -;
DR EMBL; AE017334; AAP29080.1; -;
DR EMBL; AB017225; AAT52344.1; -;
DR TIGR; BA0002; -;
DR GO; GO:0008408; F.3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F.DNA binding; IEA.
DR GO; GO:0003887; F.DNA directed DNA polymerase activity; IEA.
DR GO; GO:0008260; P.DNA replication; IEA.
DR InterPro; IPR001001; DNA polIII beta.
DR Pfam; PF00712; DNA pol3_beta; 1.
DR Pfam; PF02767; DNA pol3_beta_2; 1.
DR Pfam; PF02768; DNA pol3_beta_3; 1.
DR SMART; SM00480; POL3Bc; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome.
SQ SEQUENCE 379 AA; 42650 MW; C49D709DDBA03452E CRC64;

Query Match 37.1%; Score 692; DB 2; Length 379;
Best Local Similarity 38.8%; Pred. No. 1e-37;
Matches 148; Conservative 84; Mismatches 141; Indels 8; Gaps 7;

QY 2 IQFSINRILFIHALNTTKRAISTKNAIPILSIKIEVTSTGTTLTGSNGQISIENTIPVS 61


```
Db 1 MRFSIQDYLVRSVDVNMKAVSFRTTIPILTGKIKVAVTEGVTLTGSDADISIESFIPVE 60
Qy 62 NENAGLL-ITSPGAILLEASFPINISSLPDISINVKIEHQVQV-LTSGKSEITLKGKD 119
Db 61 EDGKEIVEVKQSGSVLQAKYFSEIVKLPKETVEI-SVENHLMTKITSGKSEFNGLND 119
Qy 120 VDQYPRLOEVSTENPLIKTLKLSIIAETAPASLQSRPILITGVHVLNHNKDFKAVA 179
Db 120 SAEYPLLPQIEBHHVFKIPTDLLKHMIRQTVFAVSTSETRPILTGVNWKVYN-SELTGIA 178
Qy 180 TDSHMSORLITLDNTSADLM--VVLPSKSLRERSAVFTDDTETVEVFFSPQILFRSEH 237
Db 179 TDSHRLAURKAKIEGIVDFQANVVPKGSNLKSLDSEEMVDIVITEYQVLFRTKH 238
Qy 238 ISFYTRLLEGNYPDTDRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLE-ITQ 296
Db 239 LLFPSRLLEGNYPDTTRLIPAESKTDIFVNTKEFLQADRSLARDGRNVVKLSTLEQ 298
Qy 297 NHISAHVNSPEVKKNEDLDIVSQGSDLTISFNPTYLIESLKAKSTVKIHFSPVRP 356
Db 299 AMLEISSNSPEIGKVEEVCQEKVDBELKISFSKAYMMMDALKALDSTEIKISFTGMRP 358
Qy 357 FTLTPGDEESFIQILTPVRT 377
Db 359 FLIRTVN-DESIILQILPVRT 378

RESULT 31
AAT29080
AC AAT29080 PRELIMINARY; PRT; 379 AA.
DT 01-JUN-2004 (TREMELrel. 27, Created)
DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit.
GN DNAN-1 OR GBAA0002.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017334; AAT29080.1; -- C49D709DBA03452E CRC64;
SQ SEQUENCE 379 AA; 42650 MW; 42650 MW; 42650 MW;

Query Match 37.1%; Score 692; DB 2; Length 379;
Best Local Similarity 38.8%; Pred. No. 1e-37;
Matches 148; Conservative 84; Mismatches 141; Indels 8; Gaps 7;

Qy 2 IOPSINRFLFIHALMTTKRAISTKNAIPILSIKIEVSTGTVLTGSGNQISIENTIPVS 61
Db 1 MRFSIQDYLVRSVDVNMKAVSFRTTIPILTGKIKVAVTEGVTLTGSDADISIESFIPVE 60
Qy 62 NENAGLL-ITSPGAILLEASFPINISSLPDISINVKIEHQVQV-LTSGKSEITLKGKD 119
Db 61 EDGKEIVEVKQSGSVLQAKYFSEIVKLPKETVEI-SVENHLMTKITSGKSEFNGLND 119
Qy 120 VDQYPRLOEVSTENPLIKTLKLSIIAETAPASLQSRPILITGVHVLNHNKDFKAVA 179
Db 120 SAEYPLLPQIEBHHVFKIPTDLLKHMIRQTVFAVSTSETRPILTGVNWKVYN-SELTGIA 178
Qy 180 TDSHMSORLITLDNTSADLM--VVLPSKSLRERSAVFTDDTETVEVFFSPQILFRSEH 237
Db 179 TDSHRLAURKAKIEGIVDFQANVVPKGSNLKSLDSEEMVDIVITEYQVLFRTKH 238
Qy 238 ISFYTRLLEGNYPDTDRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLE-ITQ 296
Db 239 LLFPSRLLEGNYPDTTRLIPAESKTDIFVNTKEFLQADRSLARDGRNVVKLSTLEQ 298
Qy 297 NHISAHVNSPEVKKNEDLDIVSQGSDLTISFNPTYLIESLKAKSTVKIHFSPVRP 356
Db 299 AMLEISSNSPEIGKVEEVCQEKVDBELKISFSKAYMMMDALKALDSTEIKISFTGMRP 358
Qy 357 FTLTPGDEESFIQILTPVRT 377
Db 359 FLIRTVN-DESIILQILPVRT 378

RESULT 31
AAT29080
AC AAT29080 PRELIMINARY; PRT; 379 AA.
DT 01-JUN-2004 (TREMELrel. 27, Created)
DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit.
GN DNAN-1 OR GBAA0002.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017334; AAT29080.1; -- C49D709DBA03452E CRC64;
SQ SEQUENCE 379 AA; 42650 MW; 42650 MW; 42650 MW;

Query Match 37.1%; Score 692; DB 2; Length 379;
Best Local Similarity 38.8%; Pred. No. 1e-37;
Matches 148; Conservative 84; Mismatches 141; Indels 8; Gaps 7;

Qy 2 IOPSINRFLFIHALMTTKRAISTKNAIPILSIKIEVSTGTVLTGSGNQISIENTIPVS 61
Db 1 MRFSIQDYLVRSVDVNMKAVSFRTTIPILTGKIKVAVTEGVTLTGSDADISIESFIPVE 60
Qy 62 NENAGLL-ITSPGAILLEASFPINISSLPDISINVKIEHQVQV-LTSGKSEITLKGKD 119
Db 61 EDGKEIVEVKQSGSVLQAKYFSEIVKLPKETVEI-SVENHLMTKITSGKSEFNGLND 119
Qy 120 VDQYPRLOEVSTENPLIKTLKLSIIAETAPASLQSRPILITGVHVLNHNKDFKAVA 179
Db 120 SAEYPLLPQIEBHHVFKIPTDLLKHMIRQTVFAVSTSETRPILTGVNWKVYN-SELTGIA 178
Qy 180 TDSHMSORLITLDNTSADLM--VVLPSKSLRERSAVFTDDTETVEVFFSPQILFRSEH 237
Db 179 TDSHRLAURKAKIEGIVDFQANVVPKGSNLKSLDSEEMVDIVITEYQVLFRTKH 238
Qy 238 ISFYTRLLEGNYPDTDRLLMTTEFEVVFNTQSLRHAMERAPLISNATQNGTVKLE-ITQ 296
Db 239 LLFPSRLLEGNYPDTTRLIPAESKTDIFVNTKEFLQADRSLARDGRNVVKLSTLEQ 298
Qy 297 NHISAHVNSPEVKKNEDLDIVSQGSDLTISFNPTYLIESLKAKSTVKIHFSPVRP 356
Db 299 AMLEISSNSPEIGKVEEVCQEKVDBELKISFSKAYMMMDALKALDSTEIKISFTGMRP 358
Qy 357 FTLTPGDEESFIQILTPVRT 377
Db 359 FLIRTVN-DESIILQILPVRT 378

RESULT 31
AAT29080
AC AAT29080 PRELIMINARY; PRT; 379 AA.
DT 01-JUN-2004 (TREMELrel. 27, Created)
DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit.
GN DNAN-1 OR GBAA0002.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017334; AAT29080.1; -- C49D709DBA03452E CRC64;
SQ SEQUENCE 379 AA; 42650 MW; 42650 MW; 42650 MW;
```

```
Db 239 LLFPSRLLEGNYPDTTRLIPAESKTDIFVNTKEFLQADRSLARDGRNVVKLSTLEQ 298
Qy 297 NHISAHVNSPEVKKNEDLDIVSQGSDLTISFNPTYLIESLKAKSTVKIHFSPVRP 356
Db 299 AMLEISSNSPEIGKVEEVCQEKVDBELKISFSKAYMMMDALKALDSTEIKISFTGMRP 358
Qy 357 FTLTPGDEESFIQILTPVRT 377
Db 359 FLIRTVN-DESIILQILPVRT 378

RESULT 32
DP3B_BACSU
ID DP3B_BACSU STANDARD; PRT; 378 AA.
AC P05649; P11571;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN Name=dnaN; Synonyms=dnaG; Ordered locus names=BSU00020;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85215612; PubMed=2987847;
RA Moriya S., Ogasawara N., Yoshikawa H.;
RT "Structure and function of the region of the replication origin of the
RT Bacillus subtilis chromosome. III. Nucleotide sequence of some 10,000
RT base pairs in the origin region.";
RL Nucleic Acids Res. 13:2251-2265(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enrian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holeappel S., Hosono S., Hulio M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Pressecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seter S.J., Serron P.,
RA Shin B.S., Soldo B., Sorokin A., Taccconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassartotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
```

```

RT subtlis.";
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE OF 1-29 FROM N.A.
RC STRAIN=168 / PSL1; Med=2846289;
RX MEDLINE=89030659; PubMed=2846289;
RA Moriya S., Fukuoka T., Ogasawara N., Yoshikawa H.;
RT Regulation of initiation of the chromosomal replication by DnaA-boxes
RT in the origin region of the Bacillus subtilis chromosome.";
RL EMBO J. 7:2911-2917(1988).
CC
CC -!- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC INDEPENDENT) along duplex DNA.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC
CC -!- SUBUNIT: Contains a core (composed of alpha, epsilon, and theta
CC chains) that can repair short gaps created by nuclease in duplex
CC DNA. For efficient replication of the long, single-stranded
CC templates, pol III requires the auxiliary chains beta, gamma, and
CC delta.
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02369; CAA26218.1; -
DR EMBL; D26185; BAA05238.1; -
DR EMBL; Z99104; CAB11778.1; -
DR EMBL; J12779; CAA31271.1; -
DR PIR; B22930; B22930.
DR Subtlis; BGI0066; dnaN.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta.1.
DR Pfam; PF02767; DNA_pol3_beta.2; 1.
DR Pfam; PF02768; DNA_pol3_beta.3; 1.
DR SMART; SM00480; POL3BC; 1.
DR TIGRFAMs; TIGR00663; dnaN; 1.
DR Complete proteome; DNA replication; DNA-directed DNA polymerase;
KW Transference.
KW
SQ SEQUENCE 378 AA; 42103 MW; C1F040B5D53026C8 CRC64;

Query Match 36.58; Score 680.5; DB 1; Length 378;
Best Local Similarity 37.4%; Pred. No. 66-37;
Matches 141; Conservative 90; Mismatches 142; Indels 7; Gaps 6;

QY 2 IQFSINRLPIHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTCNSGOISTENTIPVS 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MKFTIQKDRIVESQVDVLKAVSRTIPILTGKIVASDDGVSFTGSDSDISIESIPKE 60
QY 62 NENAGLL-ITSPGAILLEASFNNIISLPDISINKEIQHOQVLTSGKSEITLKGROV 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 EGDKEIVITEQPGSIVLOARFFSEIVKVLPMATVEIEVQNYLTIRSGKAEFNLGLDA 120
QY 121 DQYPRIQEVSTENPLIKTKLSIIAETAFASLOESRPLITGVHIVLSNHKDFKAVAT 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 DEYPLHPQIEEHAIQIPDTLLKNLRITQVAVSTSETPILTGIVNWKV-ESELLCTAT 179
QY 181 DSHRMSQRILTD-NTSADLMVVLPSKSLREFSAVFTDDIETVEVFPSPQILFRSEHIS 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 DSHRLALRAKRLDIPEDRSYNNVVPKSLTSLSKILDDNQELVDIVITETQVLFRAKNVL 239
QY 240 FYTRLLEGVYDPDRLLMTPEFETVEVFNTPQSLRHAMERAFILSNATQNGTVKLEI-TQNH 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 FSRLLDGNYPDTSLIPQDSKTEIIVTKFEQLAIDRASLRAGEGRNVNVLKAPAES 299

```

QY	299	ISAHVNSPEVKVKNEDLDIVSQSGSDLTISFNPTVYLISLKAIKSETVKIHLFSVPRPT	350
DB	300	IEISSNSPEIGKVAVIADQTEGBELNISPSPKVMLDALKVLEGAIRVSTGAMRPPL	359
QY	359	L-TPGDEESFIQLITPVRT	377
DB	360	IRTFND--ETIVQLILPVRT	377
RESULT 33			
ID	Q81JD4	PRELIMINARY;	PRT; 381 AA.
AC	Q81JD4;		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	DNA polymerase III, beta chain (EC 2.7.7.7).		
GN	ORFNames=BC0002;		
OS	Bacillus cereus (strain ATCC 14579 / DSM 31).		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=226900;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;		
RA	Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,		
RA	Kapatral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,		
RA	Chu L., Mazur M., Goltseman E., Larsen N., D'Souza M., Walunas T.,		
RA	Grechkin Y., Fusch G., Haseil Korn R., Fonstein M., Ehrlich S.D.,		
RA	Overbeek R., Kyrpides N.C.;		
RT	"Genome sequence of Bacillus cereus and comparative analysis with		
RT	Bacillus anthracis.";		
RL	Nature 423:87-91(2003).		
DR	EMBL; AE016998; AAP07112.1; -		
DR	GO; GO:0008408; F:3'-5' exonuclease activity; IEA.		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0006260; P:DNA replication; IEA.		
DR	InterPro; IPR001001; DNA_polIII_beta.		
DR	Pfam; PF00712; DNA_pol3_beta_1-		
DR	Pfam; PF02767; DNA_pol3_beta_2; 1.		
DR	Pfam; PF02768; DNA_pol3_beta_3; 1.		
DR	SMART; SMO0480; POL3bc; 1.		
DR	TIGRFAMs; TIGR00663; dnan; 1.		
KW	Nucleotidyltransferase; Transferrase.		
SEQ	SEQUENCE 381 AA; 42830 MW; 88F88D4710671163 CRC64;		
Query Match 36.5%; Score 680; DB 2; Length 381;			
Best Local Similarity 38.4%; Pred. No. 6.5e-3/;			
Matches 147; Conservative 84; Mismatches 142; Indels 10; Gaps 8;			
QY	2	IQFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSVTGLTSGNSOISTENTIPVS	61
DB	1	MRFTIQKYLVRGVQDVMKAVSSRTTIPLTGKVVATEEGVLTGSDADISTESFIPVE	60
QY	62	ENAGLL-ITSPGAILLEASFNNISLDPISINVKIEIQH-QVVLTSKSEITLKGKD	119
DB	61	ENGKEIVEIAQSGSIVLOAKYFSEIVKPLPKQTVET-SVENHFWTKIKSGSEFNGLD	119
QY	120	VDQYPRLOEVSTENPLILKTLKSIITAETAFASIQESRPILTGVHIVLSNKHDKFAVA	179
DB	120	SAEYPLLPQIEHHVFKIPDLLKHMIRQTVFAVSTSETRPILTVGNWKYN-SELTGIA	178
QY	180	TDSHRMSQRLITLD-NTSAD---LMVLPKSKSUREFSAVPTDDIETVEVFPSPQILFRS	235
DB	179	TDSHRLALRAKAKTEGNIADADEFQANVVIQKSNLSKILDESEEMVDIVITEQVQLFRT	238
QY	236	EHISFVTRLLEGNYPDTDLMLTTEFEVTVFVNTQSLRHAMERAFILSNATONGTVKLE-I	294
DB	239	KHLFFSRLLLEGNYPDTTRLIPAESKTDIFVNTKEFLQALDRASLRADGRNNVVKLSTL	298
QY	295	TQNHISAHVNSPEVKVKNEDLDIVSQSGDLTISFNPTVYLISLKAIKSETVKIHLFSV	354

```

AC QRCAL;
AD
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
DN Names=nan; OrderedLocusNames=BH0002;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=99356711; PubMed=10427704;
RA Takami H., Masui N., Nakasone K., Horikoshi K.;
RT "Replication origin region of the chromosome of alkaliphilic Bacillus
RL halodurans C-125.";
RT Biosci. Biotechnol. Biochem. 63:1134-1137(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
FUJII F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horiokoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis.";
RT Nucleic Acids Res. 28:4317-4331(2000).
CC -i- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC The beta chain is required for initiation of replication once it
CC is clamped onto DNA, it slides freely (bidirectional and ATP-
CC independent) along duplex DNA (By similarity).
CC -i- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -i- SUBUNIT: Contains a core (composed of alpha, epsilon, and theta
CC chains) that can repair short gaps created by nuclease in duplex
CC DNA. For efficient replication of the long, single-stranded
CC templates, pol III requires the auxiliary chains beta, gamma, and
CC delta (By similarity).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch.
CC
CC EMBL; AB013492; BAA82686.1; -
CC EMBL; AP001507; BAB03721.1; -
CC PIR; B83650; B83650.
CC InterPro; IPR001001; DNA_polIII_beta.
CC Pfam; PF00712; DNA_pol3_beta.1.
CC Pfam; PF02767; DNA_pol3_beta.2.1.
CC Pfam; PF02768; DNA_pol3_beta_3.1.
CC SMART; SM00480; POL3BC; 1.
CC TIGRFAMs; TIGR00663; dnan; 1.
KW Complete proteome; DNA replication; DNA-directed DNA polymerase;
KW Transference.
SQ SEQUENCE 380 AA; 42382 MW; CA72B4944E115813 CRC64;
Query Match 33.5%; Score 624.5; DB 1; Length 380;
Best Local Similarity 35.2%; Pred. No. 3.1e-33;
Matches 134; Conservative 85; Mismatches 155; Indels 7; Gaps 5
QY 2 IQFSINTEFLHALNTTKRAITSKNAPILSSIKIEVSTGVTLTGSNGQISIENTIPV- 60
DB 1 MHFVIDRDFQVNVHVKAVSSRTPIILTGIKIVADHEGVTLTGSNSDLSIEFTIPLE 60
QY 61 SNENAGLLITSPGAILLEASPFINIISSLPDISINVKIEQHVLTSGSKSITLTKGDV 120

```

Db 61 EGRQNVKQEGSVLQAKVFAEIVKVLPEQIEIHQVDSFVTTIRSGSVNENGLDP 120
 QY 121 DQYPRLOEVSTENPLILKTKLAKSIIAETAFASLOESRPLITGVHIVLNHDKFQAVAT 180
 Db 121 DEYPRLPVLEEDHVFRLPQKILKIIIRQTVFAVSTQETRPVLTGVNFEED-GILTCTAT 179
 QY 181 DSHRMSQRLITLNTSADLM--VVLPSKSLRFSFAVTDDETVEVFFSPSQILFRSH 237
 Db 180 DSHRLAMRKVPVEKNDELQFSNVVPGKSLNELSKILDENBELLDIVVTENQTLFKLKN 239
 QY 238 ISFYTLLEGNYPDTRLLMTETETEVNTQSLRHAMERAFILSNATQNTVKLEI-TQNH 299
 Db 240 MLFFSRLLEGKYEVTKMIPKKAETSAFVTKAFLQTLERALLREGKNQVINKTLCD 299
 QY 297 NHTSAHVNSPEVGVKNEDLDIVSQSGDLTISFNPVLYLESKAIKSETVKIHLSPVRP 356
 Db 300 GVVEVTAITPEIGKVTENATQGLEGEELRISFNGKNVIDALKVDSSEIHIAFTGAMSP 359
 QY 357 FYLTPGDEESFIQLITPVRT 377
 Db 360 FVLSPDTHDQS-LHLFSPVRT 379

RESULT 36

Q6HI40 PRELIMINARY; PRT; 376 AA.
 AC Q6HI40;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DE DNA polymerase III, beta subunit (EC 2.7.7.7).
 GN Name=dnaN; ORNAMES=BT927.2460;
 OS Bacillus thuringiensis serovar konkukian str. 97-27.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
 OC Bacillus thuringiensis serovar konkukian.
 OX NCBI_TaxID=281309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017355; AAT61374.1; --
 DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta; 1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3bc; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 KW Nucleotidyltransferase; Transferase.
 SQ SEQUENCE 376 AA; 41592 MW; E098AA29FA486A42 CRC64;

Query Match 30.6%; Score 569.5; DB 2; Length 376;
 Best Local Similarity 34.4%; Pred. No. 1.3e-29;
 Matches 130; Conservative 80; Mismatches 161; Indels 7; Gaps 7;
 QY 2 IQFSINRTLFHALNTHKRAISTKNAIPILSSIKIEVSTGTGVLTSNGOISIENTIP-V 60
 Db 1 MEFIIVNHKHTQALSEVSKAISTKAVIPILSGIKITADQSGITLIASNSNIFIEKIPSV 60
 QY 61 SNEAGLLITSPGAILLEASFINISSLDPISINVKIEHQHVLTSGKSEITLKQDV 120
 Db 61 IDDSQITLQAGTIVPAKFIIEIKKMP5-DIVIKSKNEQITIOSGITLNLNGFPA 119
 QY 121 DQYPRLOEVSTENPLILKTKLAKSIIAETAFASLOESRPLITGVHIVLNHDKFQAVAT 180
 Db 120 NEFENVPLQDGHTEIQIETKQLIDAFKQTVFAVAKNESRPLVIGVHIEL-DHKLICAA 178
 QY 181 DSHRMSQRLITLNTSADLMVLPKSLRFSFAVTDDETVEVFFSPSQILFRSHISF 240
 Db 179 DSHRLAIR-ETLISNTNMKNCIVPESATINELLKLMNSNLEFVSILYESHIEFTFGTTL 237

QY 241 YTRLEGNYPDTRLLMTETETEVNTQSLRHAMERAFILSNATQNTVKLEI-TQNH 299
 Db 238 YSRLEGEKYPNISTLIPNEFQTVINIDRQRMLOGVDRSSLLASEWANNVNLIVNESTI 297
 QY 300 SAHVNSPEVGVKNEDLDIVSQSG-DLTISFNPVLYLESKAIKSETVKIHLSPVRPPT 358
 Db 298 QISSNASQIGIKSETQIDIGLQGEKQNLISFDGRFMVEVLKAIKEETITLISFGGSMRPL 357
 QY 359 LTPGDEESFIQLITPVR 376
 Db 358 IEAG-EQSAVVHLISPVR 374
 RESULT 37
 Q6IPW0 PRELIMINARY; PRT; 376 AA.
 AC Q6IPW0; Q6HY29; Q6KS44;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE DNA polymerase III, beta subunit.
 GN Name=dnaN-2; OrderedLocusNames=BA2684, BAS2499; ORNAMES=GBAA2684;
 OS Bacillus anthracis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate Porton;
 RX MEDLINE=22608414; PubMed=1271629; DOI=10.1038/nature01586;
 RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
 RA Holtzapfel E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Cline R.T., Redmond C., Thwaitte J.E., White O.,
 RA Salzberg S.L., Thomas B., Friedlander A.M., Kohler T.M.,
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
 RA "The genome sequence of Bacillus anthracis Ames and comparison to
 RT closely related bacteria."
 RL Nature 423:81-86(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate 0581;
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
 RA Fraser C.M.;
 RT "Bacillus anthracis comparative genomics";
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Stearne;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017032; AAP26527.1; --
 DR EMBL; AE017334; AAT31800.1; --
 DR EMBL; AE017225; AAT54810.1; --
 DR TIGR; BA2684; --
 DR GO; GO:0008408; F:3'-5', exonuclease activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004260; P:DNA replication; IEA.
 DR InterPro; IPR001001; DNA_polIII_beta.
 DR Pfam; PF00712; DNA_pol3_beta; 1.
 DR Pfam; PF02767; DNA_pol3_beta_2; 1.
 DR Pfam; PF02768; DNA_pol3_beta_3; 1.
 DR SMART; SM00480; POL3bc; 1.
 DR TIGRFAMs; TIGR00663; dnan; 1.
 SQ SEQUENCE 376 AA; 41705 MW; E57D297F75E96F4E CRC64;

```
Query Match      29.9%; Score 557.5; DB 2; Length 376;
Best Local Similarity 33.9%; Pred. No. 8.4e-29;
Matches 128; Conservative 83; Mismatches 160; Indels 7; Gaps 7;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVSTGTGTLTGSNGQISIENTIPVS 61
DB 1 MEFIVNHKHFQALSDVSKAISTKAIIPILSGIKITADQSGITLIASNSNPIEFKFIPL 60
QY 62 NENAGL-LITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
DB 61 IDEDEQITILQAGTIVVPKAFYFIEIIKKWPS-DIVIKSKNEQITIIQSGEITLNLGPPA 119
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFAASLQESRPILTGVHIVLSNHKOPKAVAT 180
DB 120 NEFFNVPQIDHTEIQITKQIDAFKQTVFAVAKNESRHLVTGVHIEL-DHNKLI CAAT 178
QY 181 DSHRMSORLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
DB 179 DSHRLAIR-ETLISNKMKANCIVPSATINELLKLMNSNLEFVSIYLSSEHIIFTGTFTL 237
QY 241 YTRLLEGNYPTDRLMTTEFEVFNFTQSLRHAMERAFILSNATQNGTVKLEI-TQNH 299
DB 238 YSRLIEGKYPNISTLIPIEFQTVINIDRQMLQGVDRSSLLASEWANNVNLVINESTI 297
QY 300 SAHVNSPEVGVKNEDLDI-VSQSGSDLTISFNPTVLYIESLKAISKSETVKIHLSPVRPFT 358
DB 298 QISSNASQIGKISEKQIDVIGKKQLNISPGRFMDALRAIKETVTLTSFGSGMRPIL 357
QY 359 LTPGDEEESFQILTPVR 376
DB 358 IEAGTQSAI-IHLISPV 374

RESULT 38
AAT31800
ID AAT31800 PRELIMINARY; PRT; 376 AA.
AC AAT31800;
DT 01-JUN-2004 (TREMBLrel. 27, Created)
DT 01-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit.
GN DNAN-2 OR GBAA2684.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RA "Bacillus anthracis comparative genomics.";
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE017334; AAT31800.1; -.
DR SEQUENCE 376 AA; 41705 MW; E57D297F75E96F4E CRC64;

Query Match      29.9%; Score 557.5; DB 2; Length 376;
Best Local Similarity 33.9%; Pred. No. 8.4e-29;
Matches 128; Conservative 83; Mismatches 160; Indels 7; Gaps 7;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVSTGTGTLTGSNGQISIENTIPVS 61
DB 1 MEFIVNHKHFQALSDVSKAISTKAIIPILSGIKITADQSGITLIASNSNPIEFKFIPL 60
QY 62 NENAGL-LITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
DB 61 IDEDEQITILQAGTIVVPKAFYFIEIIKKWPS-DIVIKSKNEQITIIQSGEITLNLGPPA 119
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFAASLQESRPILTGVHIVLSNHKOPKAVAT 180
DB 120 NEFFNVPQIDHTEIQITKQIDAFKQTVFAVAKNESRHLVTGVHIEL-DHNKLI CAAT 178
QY 181 DSHRMSORLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
DB 179 DSHRLAIR-ETLISNKMKANCIVPSATINELLKLMNSNLEFVSIYLSSEHIIFTGTFTL 237
QY 241 YTRLLEGNYPTDRLMTTEFEVFNFTQSLRHAMERAFILSNATQNGTVKLEI-TQNH 299
DB 238 YSRLIEGKYPNISTLIPIEFQTVINIDRQMLQGVDRSSLLASEWANNVNLVINESTI 297
QY 300 SAHVNSPEVGVKNEDLDI-VSQSGSDLTISFNPTVLYIESLKAISKSETVKIHLSPVRPFT 358
DB 298 QISSNASQIGKISEKQIDVIGKKQLNISPGRFMDALRAIKETVTLTSFGSGMRPIL 357
QY 359 LTPGDEEESFQILTPVR 376
DB 358 IEAGTQSAI-IHLISPV 374

RESULT 38
AAT31800
ID AAT31800 PRELIMINARY; PRT; 376 AA.
AC AAT31800;
DT 01-JUN-2004 (TREMBLrel. 27, Created)
DT 01-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit.
GN DNAN-2 OR GBAA2684.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RA "Bacillus anthracis comparative genomics.";
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE017334; AAT31800.1; -.
DR SEQUENCE 376 AA; 41705 MW; E57D297F75E96F4E CRC64;

Query Match      29.7%; Score 553.5; DB 2; Length 376;
Best Local Similarity 33.9%; Pred. No. 1.5e-28;
Matches 128; Conservative 85; Mismatches 158; Indels 7; Gaps 7;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVSTGTGTLTGSNGQISIENTIPVS 61
DB 1 MEFIVNHKHFQALSDVSKAISTKAIIPILSGIKITADQSGITLIASNSNPIEFKFIPL 60
QY 62 NENAGL-LITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
DB 61 IDEDEQITILQAGTIVVPKAFYFIEIIKKWPS-DIVIKSKNEQITIIQSGEITLNLGPPA 119
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFAASLQESRPILTGVHIVLSNHKOPKAVAT 180
DB 120 NEFFNVPQIDHTEIQITKQIDAFKQTVFAVAKNESRHLVTGVHIEL-SNNK-LICAAT 178
QY 181 DSHRMSORLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
DB 179 DSHRLAIR-ETLISNKMKANCIVPSATINELLKLMNSNLEFVSIYLSSEHIIFTGTFTL 237
QY 241 YTRLLEGNYPTDRLMTTEFEVFNFTQSLRHAMERAFILSNATQNGTVKLEI-TQNH 299
DB 238 YSRLIEGKYPNISTLIPIEFQTVINIDRQMLQGVDRSSLLASEWANNVNLVINESTI 297
QY 300 SAHVNSPEVGVKNEDLDI-VSQSGSDLTISFNPTVLYIESLKAISKSETVKIHLSPVRPFT 358
DB 298 QISSNASQIGKISEKQIDVIGKKQLNISPGRFMDALRAIKETVTLTSFGSGMRPIL 357
QY 359 LTPGDEEESFQILTPVR 376
DB 358 IEAGTQSAI-IHLISPV 374

RESULT 39
Q81CQ4
ID Q81CQ4 PRELIMINARY; PRT; 376 AA.
AC Q81CQ4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE DNA polymerase III, beta chain (EC 2.7.7.7).
GN ORENAMES-BC2693;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapatal V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Hasekorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:87-91(2003).
DR EMBL; AE017006; AAP09649.1; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR01001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta_1.
DR Pfam; PF02767; DNA_pol3_beta_2_1.
DR Pfam; PF02768; DNA_pol3_beta_3_1.
DR SMART; SM00480; POL3BC; 1.
DR TIGRFAMs; TIGR00663; dnan; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 376 AA; 41736 MW; 9E9DF4C4AD18623B CRC64;

Query Match      29.7%; Score 553.5; DB 2; Length 376;
Best Local Similarity 33.9%; Pred. No. 1.5e-28;
Matches 128; Conservative 85; Mismatches 158; Indels 7; Gaps 7;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVSTGTGTLTGSNGQISIENTIPVS 61
DB 1 MEFIVNHKHFQALSDVSKAISTKAIIPILSGIKITADQSGITLIASNSNPIEFKFIPL 60
QY 62 NENAGL-LITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
DB 61 IDEDEQITILQAGTIVVPKAFYFIEIIKKWPS-DIVIKSKNEQITIIQSGEITLNLGPPA 119
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFAASLQESRPILTGVHIVLSNHKOPKAVAT 180
DB 120 NEFFNVPQIDHTEIQITKQIDAFKQTVFAVAKNESRHLVTGVHIEL-SNNK-LICAAT 178
QY 181 DSHRMSORLITLNTSADLMVLPKSLRSESAVFTDDIETVEVFFSPQILFRSEHISF 240
DB 179 DSHRLAIR-ETLISNKMKANCIVPSATINELLKLMNSNLEFVSIYLSSEHIIFTGTFTL 237
QY 241 YTRLLEGNYPTDRLMTTEFEVFNFTQSLRHAMERAFILSNATQNGTVKLEI-TQNH 299
DB 238 YSRLIEGKYPNISTLIPIEFQTVINIDRQMLQGVDRSSLLASEWANNVNLVINESTI 297
QY 300 SAHVNSPEVGVKNEDLDI-VSQSGSDLTISFNPTVLYIESLKAISKSETVKIHLSPVRPFT 358
DB 298 QISSNASQIGKISEKQIDVIGKKQLNISPGRFMDALRAIKETVTLTSFGSGMRPIL 357
QY 359 LTPGDEEESFQILTPVR 376
DB 358 IEAGTQSAI-IHLISPV 374
```

Db 179 DSHRLAIR-ETLLSDVKANCIVPSATINELLKMNSEFVYIYFSESHIIFTFTGTTTL 237
QY 241 YTRLLEGNYPDTRLMTETETEVNTQSLRHAMERAFILSNATQNGTVKLE-ITQNH 299
Db 238 YSLIEGKYPNINLIPDNFKTIINVRKILQGVDRSSLASEWANNVNLINVESTI 297
QY 300 SAHVNSPEVGVKNEDLDI-VSQSGDLTISFNPTYLIESLKAIKSETVKIHLSPVRPPT 358
Db 298 KISSNASQIGKISSETQOIDAIOEKQLNISFDGRFMDALRAIKETITLSFGGSMRPIL 357
QY 359 LTPGDEEESFIQLITPVR 376
Db 358 IEAG-EQSAAVHLISPV 374

RESULT 40

Q737DO PRELIMINARY; PRT; 376 AA.
AC Q737DO;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE DNA polymerase III, beta subunit (EC 2.7.7.7).
GN Name-dnaN; OrderedlocusNames=BCE2720;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
adaptations and a large plasmid related to Bacillus anthracis pX01.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017272; AAS41632.1; -;
DR TIGR; BCE2720; -;
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001001; DNA polIII beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02767; DNA_pol3_beta; 2.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRfams; TIGR00663; dnan; 1.
SQ Complete proteome; Nucleotidyltransferase; Transferase.
KW SEQUENCE 376 AA; 41801 MW; ED9AFBCF7E5CBE31 CRC64;

Query Match 29.0%; Score 539.5; DB 2; Length 376;
Best Local Similarity 33.3%; Pred. NO. 1.3e-27;
Matches 126; Conservative 82; Mismatches 163; Indels 7; Gaps 7;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSKIEVTSTGTTLTSGNGQISIENTIPVS 61
Db 1 MEFIWNKHGFTQALSEVSKALSTKAFIPILSGIKITANQSGITLIANSNIFIEKFISS 60
QY 62 NENAGLLIT-SFGAILLEASFMINISSLPDISINVKEIQHVVLTSKSEITLKGKV 120
Db 61 MDDEQITTTLQAGTIVPAKYFIBIKKMP-S-DIVIKSKNEQITIQSEETLNLGPPA 119
QY 121 DQYPLQEVSTENPLIKTLKILIAETAFAASLQESRRPILTCGHVILSNHDKPKAVAT 180
Db 120 SEFPNVPQIEHAEIQETKQIDAFKQTVFAVAKNESRPVLTGVHIELDYK-LICAA 178
QY 181 DSHRMSQRLITLDNTADLMVLPKSLRPFSAVFTDDIETVEVFFSPQILFRSEHISF 240
Db 179 DSHRLAIR-ETQISTNNKANCIVPSATINELLKMNSEFVYIYFSESHIIFTFTGTT 237
QY 241 YTRLLEGNYPDTRLMTETETEVNTQSLRHAMERAFILSNATQNGTVKLEI-TQNH 299
Db 238 YSLIEGKYPNINLIPDNFKTIINVRKILQGVDRSSLASEWANNVNLINVESTI 297

QY 300 SAHVNSPEVGVKNEDLDIVSQSG-DLTISFNPTYLIESLKAIKSETVKIHLSPVRPPT 358
Db 298 KVSSNASQIGKIYETQQVDAIHGEKQLHISFDGRFMDALRAIKETITLSFGGSMRPIL 357
QY 359 LTPGDEEESFIQLITPVR 376
Db 358 IEAGTQSA-ITLISPV 374

Search completed: January 28, 2005, 16:59:36
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2005, 16:50:32 ; Search time 41 Seconds
(without alignments)
887.072 Million cell updates/sec

Title: US-10-048-071-28
Perfect score: 1863
Sequence: 1 MIQFSINRTLFTHALNTTKR.....LTPGDEBESFIQLITPVRTN 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1389	74.6	378	2 B95000	DNA polymerase III
2	1384	74.3	378	2 B97872	probable DNA-direc
3	1052	56.5	380	2 B86625	DNA-directed DNA p
4	1052	56.5	380	2 T30306	DNA-directed DNA p
5	730.5	39.2	377	1 S54708	DNA-directed DNA p
6	710	38.1	381	2 AC1433	DNA polymerase III
7	703	37.7	381	2 AC1432	DNA polymerase III
8	680.5	36.5	378	1 B22930	DNA-directed DNA p
9	624.5	33.5	380	2 B83650	DNA polymerase III
10	457.5	24.6	366	2 B96900	DNA polymerase III
11	423	22.7	363	1 S35733	DNA-directed DNA p
12	396.5	21.3	397	2 S70987	dnan protein - Myc
13	364	19.5	402	2 F70850	DNA-directed DNA p
14	362.5	19.5	367	1 DJPS3P	DNA-directed DNA p
15	354	19.0	366	1 A64107	DNA-directed DNA p
16	352	18.9	366	2 H82859	DNA polymerase III
17	350.5	18.8	367	2 F83644	DNA polymerase III
18	344	18.5	379	2 G97772	DNA-directed DNA p
19	339	18.2	366	2 F83276	DNA polymerase III
20	336.5	18.1	391	2 T47420	DNA-directed DNA p
21	335.5	18.0	399	2 T10002	DNA-directed DNA p
22	332	17.8	366	2 AB0958	DNA polymerase III
23	325	17.4	372	2 AE2613	DNA polymerase III
24	325	17.4	403	2 C97395	DNA polymerase III
25	323	17.3	381	2 B71700	DNA polymerase III
26	319.5	17.1	366	2 E72400	DNA polymerase III
27	318	17.1	366	1 DJBC3B	DNA-directed DNA p
28	318	17.1	366	2 D91208	DNA polymerase III
29	318	17.1	366	2 F86054	DNA polymerase III

30	317	17.0	366	2 AF0497	DNA-directed DNA p
31	315.5	16.9	456	2 AF2535	DNA polymerase III
32	314	16.9	387	2 AD2057	DNA polymerase III
33	314	16.9	397	2 AH3494	DNA-directed DNA p
34	307	16.5	367	1 JQ0734	DNA-directed DNA p
35	296.5	15.9	366	2 C81713	DNA polymerase III
36	296	15.9	376	2 B41870	DNA-directed DNA p
37	295.5	15.9	366	2 H72090	DNA polymerase III
38	295.5	15.9	366	2 B86533	DNA polymerase III
39	293.5	15.8	366	2 F81578	DNA polymerase III
40	292.5	15.7	416	2 E71559	probable DNA pol I
41	288	15.5	366	2 C84931	DNA-directed DNA p
42	287	15.4	372	2 C87268	DNA polymerase III
43	275	14.8	366	1 JC1159	DNA-directed DNA p
44	272.5	14.6	363	2 C70462	DNA polymerase III
45	270	14.5	367	2 C81030	DNA polymerase III

ALIGNMENTS

RESULT 1

B95000
DNA polymerase III, beta chain [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95000
R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heit, on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95000
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <KUR>
A:Cross-references: UNIPROT:O06672; GB:AE005672; PIDN:AAK74195.1; PID:g14971466; GSPDB: A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0002
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match	74.6%	Score 1389;	DB 2;	Length 378;
Best Local Similarity	72.2%	Pred. NO. 1.6e-83;		
Matches 273;	Conservative 50;	Mismatches 55;	Indels 0;	Gaps 0;
Qy	1	MIQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGOISIENTIPV	60	
Db	1	MIHFSINKNLFQALNTTKRAISSKNAIPILSTVKIDVTNEGTLIGSGQISIENTISQ	60	
Qy	61	SNENAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV	120	
Db	61	KNEDAGLLITSLGSLLEASFFINVVSLPDVTLDFKIEQNVILTSKSEITLKGKDS	120	
Qy	121	DQYPLQVSTENPILTKLKSIIATAFAASQESRPILTGVIHVLVSNHOKFVAT	180	
Db	121	EQYPRIQEISASTPLILETKLKIINETAFAASQESRPILTGVIHVLVSNHOKFVAT	180	
Qy	181	DSHRMSQRLITLNTSADLVVLPSKSLRESAVTDDIETVEVFPSQIILFRSEHISF	240	
Db	181	DSHRLSQKLLKLNKDDDFVWIPSRSLRESAVTDDIETVEVFPSQIILFRSEHISF	240	
Qy	241	YTRLLEGNYPDTDRLLMTTEFEVVFNTQSLRHAMERAFILSNATQNGTVKLEITQNHIS	300	
Db	241	YTRLLEGNYPDTDRLLMTTEFEVVFNTQSLRHAMERAFILSNATQNGTVKLEIKDGVVS	300	
Qy	301	AHVNSPEYGVKNEDLDIVSQSGSDLTISFNPTYLIESLKAIKSETVKIHFSPVPFTLT	360	
Db	301	AHVHSPGVKNEEDTDQVTGDLTISFNPTYLIDSLKALNSEKVTISFISAVRPFTLV	360	
Qy	361	PGDEESFIQLITPVRTN	378	

Query Match	39.2%;	Score	730.5;	DB	1;	Length	377;
Best Local Similarity	39.3%;	Pred.	No. 1.8e-40;				
Matches	149;	Conservative	90;	Mismatches	135;	Indels	5;
Gaps	4						
Qy	1	MIQFSINRNLFIHAIHNTTKRAISTKNAJLPILSSIKIEVTSTGCVTLTSGNGOISIENTIPV	60				
Db	1	MMEFTIKRDYFETQNDLTKAISPRFTUPTUIGIKIDAKEHEVILUTGSDSEISIIITPK	60				
Qy	61	SNENAGLL-ITSPGAILEASFFNIISLDPDISINVKIEIBQHQVVLSTGSKSEITLKGKD	119				
Db	61	TVQGEDIVNSETGSVLPGRFFVDI IKKLPGKDVKLSTNEQFQTLITSGHSEFNLGLD	120				
Qy	120	VDQYPRLOEVSTENPLIIKTLKLSIIAETAFASLOESRPLITTCGVHIVLSNHRKDFKAVA	179				
Db	121	PDQYPLLPQSRDDAIQUSVVLKNVIAQTFNFAVSTSETRPLVLTGNVLIOEN-ELICTA	179				
Qy	180	TDHRMSORLITLTONTSADLMVLPKSLRBFSAVFTDDIETVEVFFSPQSLPRSEHIS	239				
Db	180	TDSHRLAVRKLQLEDVSENKNVITPGKALAEINKMSDNEEDIDIFFASNQVLPKVGNNV	239				
Qy	240	PYTLLLEGNYDDTRLMTBETEVVFNQTSURHAMERAF LISNATQNGTVKLETQNHII	299				
Db	240	FISRLLEGHYDDTRLFPENYEIKLSDINGEFYHAIDRASLLAREGGNNVILKSTGDDVV	299				
Qy	300	SAHVNSPEVGKVNEDLDITVQSGSDLTISFNPTYLIESLKATKSTVKIHFSLPVRPFTL	359				
Db	300	ELSTSPETIGTVKEVDANDVEGSLKISFNKSNKMYMDALKALINDVEVEFFGTMKPFIL	359				
Qy	360	TP-GDEESFIQLITTPVRT	377				
Db	360	KPKGD--DSVTQILILPIRT	376				

RESULT 7
AC1432
DNA poly
C/Species
C/Date:
C/Access
R/Glaser
.; Domin
D.; Jone
Science
A/Author
ok, C.;
A/Title:
A/Refer

[illegible]

A:Accession: AC1432
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-381 <GLA>
A:Cross-references: UNIPROT:Q8YAW1; GB:NC_003210; PIDN:CAC98217.1; PID:gl6409361; GSPDB:
A:Experimental source: strain EGB-e
C:Genetics:
A:Gene: dnaN
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 37.7%; Score 703; DB 2; Length 381;
Best Local Similarity 38.8%; Pred. No. 1.2e-38;
Matches 148; Conservative 93; Mismatches 134; Indels 6; Gaps 5;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIP-V 60
DB 1 MKFVIERDLVQAVNEVTRAIARTIPTLTGKIIVNDEGVTLTGSDSDISIEAFPLI 60

QY 61 SHENAGLLTSGAILLEASFFINISSLPDISINKEIQHVLTGSKSEITLKGKDV 120
DB 61 ENDEVIVEVESGGIVLOSQYFGDIVRRLPEENVEIETSNYQTNISSQASFTLGLDP 120

QY 121 DOYPRLOEVSTENPILTKLKSIIAETAPAAASIQESRPILTGVHIVLSNHKDFKAVAT 180
DB 121 MEYPKLPEVDGKTIKIPINVLKIVRQTVFAVSAIEVRPVLTVGNVLIKENK-LSAVAT 179

QY 181 DSHRMSQRLITLD-NTSADLMVLPKSLRFSFAVFTDDIETVEVFPSQILFRSEHIS 239
DB 180 DSHRLALREIPLTETDIEYINIVFGKSLSELNKLDDASESIEMTLANNQILFKLDLL 239

QY 240 FYTRLLEGNYPTDRLMTEPTEVFNTQSLRHAMERAFILSNATQGTVKLEITON-H 298
DB 240 FYSLRLEGSYPDTSRLIPTDTKSELVINSKAPLQADIRASLLARENNVVKLMTLENGQ 299

QY 299 ISAHVNSPEVGVKNEDLDIVSOGSDLTISFNPTVLIIESLKAIKSETVKIHFSPVRPT 358
DB 300 VEVNSNSPEVGVNSVNSFSQSTGEEKISFNKGYMDALRAFEQDDIQISFGTMRPFV 359

QY 359 LTPGD-BEESFIQITPVRT 377
DB 360 LRPKDAANPNEILQITPVRT 380

RESULT 8
B22930
N:Alternate names: 42K protein (oriC region)
C:Species: Bacillus subtilis
C:Date: 21-May-1988 #sequence revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: B22930; S66032; I40393; A69618
R:Moriya, S.; Ogasawara, N.; Yoshikawa, H.
Nucleic Acids Res. 13, 2251-2265, 1985
A:Title: Structure and function of the region of the replication origin of the Bacillus
A:Reference number: A94702; MUID:85215612; PMID:2987847
A:Accession: B22930
A:Molecule type: DNA
A:Residues: 1-378 <OR>
A:Cross-references: UNIPROT:P05649; GB:D26185; NID:G467326; PIDN:BAA05238.1; PID:G467392
R:Ogasawara, N.; Moriya, S.; von Meyenburg, K.; Hansen, F.G.; Yoshikawa, H.
EMBO J. 4, 3345-3350, 1985
A:Title: Conservation of genes and their organization in the chromosomal replication ori
A:Reference number: A26057; MUID:86135970; PMID:3004954
A:Contents: annotation
R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom
A:Reference number: S65967; MUID:96051385; PMID:7584024
A:Accession: S66032
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <OGA>
A:Cross-references: EMBL:D26185; NID:G467326; PIDN:BAA05238.1; PID:G467392
R:Moriya, S.; Fukuoka, T.; Ogasawara, N.; Yoshikawa, H.

EMBO J. 7, 2911-2917, 1988
A:Title: Regulation of initiation of the chromosomal replication by DnaA-boxes in the ori
A:Reference number: I40391; MUID:89030659; PMID:2846289
A:Accession: I40393
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-30 <RES>
A:Cross-references: EMBL:X12779; NID:G39878; PIDN:CAA31271.1; PID:G39879
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, V.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerison, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; E
Nature 390, 249-256, 1997
A:Authors: Fulgner, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Ketter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69618
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-378 <UN>
A:Cross-references: GB:Z99104; GB:AL009126; NID:G2632267; PIDN:CAB11778.1; PID:G2632269
A:Experimental source: strain 168
C:Genetics:
A:Gene: dnaN
A:Map position: 0
C:Superfamily: DNA-directed DNA polymerase III beta chain
C:Keywords: DNA replication initiation; nucleotidyltransferase

Query Match 36.5%; Score 680.5; DB 1; Length 378;
Best Local Similarity 37.1%; Pred. No. 3.4e-37;
Matches 141; Conservative 90; Mismatches 142; Indels 7; Gaps 6;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIPVS 61
DB 1 MKFTIQDRLVESQDVLKAVSSRTTIPILTKIKIVASDDGVSTGSDSDISIESFIPKE 60

QY 62 NENAGLL-ITSPGAILLEASFFINIISLSPDISINKEIQHVLTGSKSEITLKGKDV 120
DB 61 EGDKEIVTIEQPGSIVQARFSEIVKLPMTATVEIQVQYLTIIIRSGAEFNLGLDA 120

QY 121 DOYPRLOEVSTENPILTKLKSIIAETAPAAASIQESRPILTGVHIVLSNHKDFKAVAT 180
DB 121 DEYPLHPQIEBHAIQIPTDILLKNIQRQTVFAVSTSETRPILTVGNWKV-BQSELLCTAT 179

QY 181 DSHRMSQRLITLD-NTSADLMVLPKSLRFSFAVFTDDIETVEVFPSQILFRSEHIS 239
DB 180 DSHRLALRKAKLDIPEDRSYNNVTPGKSLTSLKILDDNQLDQELDIVITETQVLKAKNVL 239

QY 240 FYTRLLEGNYPTDRLMTEPTEVFNTQSLRHAMERAFILSNATQGTVKLEI-TQNH 298
DB 240 FFSRLLDGNYPTDTSLLIPQDSKTEIIVNTKEFLQADIRASLLAREGNVNVKLSAKPAES 299

QY 299 ISAHVNSPEVGVKNEDLDIVSOGSDLTISFNPTVLIIESLKAIKSETVKIHFSPVRPT 358
DB 300 IEISSNSPEIGKVVEAIVADQIEGELNIFSPKPYMLDALKVLEGAIRVFTGAMRPF 359

QY 359 L-TPGDEEESFIQITPVRT 377
DB 360 IRTEND--ETIVQLLPVRT 377

RESULT 9
B83650
DNA polymerase III (beta subunit) dnaN [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: B96900
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: B96900
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-380 <STO>
A/Cross-references: UNIPROT:Q9RCAL; GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA0307
A/Experimental source: strain C-125
C/Genetics:
A/Gene: dnaN
C/Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 33.5%; Score 624.5; DB 2; Length 380;
Best Local Similarity 35.2%; Pred. No. 1.6e-33;
Matches 134; Conservative 85; Mismatches 155; Indels 7; Gaps 5;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPV- 60
DB 1 MHFVIDRDIQVQNVHVS KAVSSRTTIPILTGKIVADHEGVTLTGSDSDISIEFIPL 60

QY 61 SNEAGLLITSPGAILLEASFPINIISSIPDISINVKIEHQVLTSGKSEITLKGKDV 120
DB 61 EGDQRQNVKQBSIVLQAKVFAEIVKLPQEQIEIHVQDSFVTTIRSGSSVFNGLDLP 120

QY 121 DQYRLQEVSTENPLILTKLKSIIAETAFASQESRPILTGVHIVLSNHNKDKAVAT 180
DB 121 DEYPLPLVLEEDHVRFLQKILKDIIRQTVFVASTQETRPVLTGVNFIED-GILTCAT 179

QY 181 DSHRMSORLITLNDTSADLM---VVLPSKSLREFSAVFTDDIETVEFPSPQILFRSEH 237
DB 180 DSHRLAKVPVEKNDELQFSNVVPIPKSLNELSKILDENEELLDIVVTENQTLFLKLN 239

QY 238 ISFYTRLEGVYDPTDRLLMTTEFEVFNVTQSLRHAMERAPLISNATQNGTVKLE-ITQ 296
DB 240 MLFFSRLEGGYVPTKMPKPEAKTSFAVHTKAFLOTLEALLLREGKNQVNLKGLD 299

QY 297 NHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESKAIKSETVKIHFSLSPVRP 356
DB 300 GVEVTAITPEIGKTENVATQGLEGEEDRISFNKVNIDALKVDSSEIHIAFTGAWSP 359

QY 357 FTLPDGEESFIQILTPVT 377
DB 360 FVLSPTDHDQS-LHLFSPVT 379

RESULT 10
B96900
DNA polymerase III beta chain [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: B96900
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: B96900
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-366 <KUR>
A/Cross-references: UNIPROT:Q97N34; GB:AE001437; PIDN:AAK77989.1; PID:g15022819; GSPDB:C
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC0002

Query Match 24.6%; Score 457.5; DB 2; Length 366;
Best Local Similarity 29.9%; Pred. No. 1.2e-22;
Matches 112; Conservative 91; Mismatches 159; Indels 13; Gaps 7;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 61
DB 1 MKFICEKNILQEAIIATAQAVTGKSTMPVLOGILMSVQNNELTLGSDISLTIETKINVE 60

QY 62 NENAGLLITSPGAILLEASFPINIISSIPDISINVKIEHQVLTSGKSEITLKGKDV 121
DB 61 -----VLEEGKVLDARLLSEIRKLPNSKVEIQTIENNCEVITCKNSKUTLVYLPN 113

QY 122 QYPRLOEVSTENPLILTKLKSIIAETAFASQESRPILTGVHIVLSNHNKDKAVATD 181
DB 114 DPSPLEIDENSIFKINQKTLTKMIGTIFAQOETRPILTGVLFKDSK-LNLVAID 172

QY 182 SHRMSORLITLNDTSADLMVVLPSKSLREFSAVFTDDIETVEFPSPQILFRSEHSFY 241
DB 173 GYRLALRSQYIDNETS-INAVIPGKTLNEVKILEDDGD-VNITFTSNHILFNLGNTKII 230

QY 242 TRLEGVYDPTDRLLMTTEFEVFNVTQSLRHAMERAPLISNATQNGTVKLEITQNHISA 301
DB 231 SRLLEGEFIKYSIPEEYNINIVARKEELDCIERASIMAKDGNLILKLDI-EDDVM 289

QY 302 HVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESKAIKSETVKIHFSLSPVRPFTLTP 361
DB 290 ITSNSQLGNVREENIILQ-GQPLKIAFNKYLDIVLKMNQEEIWMNFSSSISPCIIKN 348

QY 362 GDEESFIQILTPVR 376
DB 349 KENDDS-TYLLTPVR 362

RESULT 11

S35733
DNA-directed DNA polymerase (EC 2.7.7.7) III beta chain - Spiroplasma citri
N/Alternate names: dnaN
C/Species: Spiroplasma citri
C/Date: 13-Jan-1995 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C/Accession: S35733
R:ife, F.; Laigret, F.; Bove, J.
Submitted to the EMBL Data Library, December 1992
A/Description: Nucleotide sequence and genetic organization at the replication origin (o
A/Cross-references: UNIPROT:P34029; EMBL:Z19108; NID:g49345; PIDN:CAA79522.1; PID:g49347
A/Reference number: S35732
A/Accession: S35733
A/Molecule type: DNA
A/Residues: 1-363 <YEP>
C/Genetics:
A/Gene: dnaN
A/Genetic code: SGC3
C/Superfamily: DNA-directed DNA polymerase III beta chain
C/Keywords: DNA replication initiation; nucleotidyltransferase

Query Match 22.7%; Score 423; DB 1; Length 363;
Best Local Similarity 27.8%; Pred. No. 2.2e-20;
Matches 104; Conservative 90; Mismatches 164; Indels 16; Gaps 7;

QY 5 SINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVSNE 64
DB 4 NIKERDILDELLKVSRIISQKTLIPSLGILIEVKDKITFTTSDGTSIKSEI-MGND- 61

QY 65 AGLITSPGAILLEASFPINIISSIPDISINVKIEHQVLTSGKSEITLKGKDVDP 124
DB 62 --LMTIRIGVLIIRKNKFIIVEIRKIEDEFTLEVVEGNLIKIKANNFOSVLTLSADYP 119

QY 125 RLQEVSTENPLILTKLKSIIAETAFASQESRPILTGVHIVL-SNHNKDKAVATDSH 183
DB 120 HLFSETEGEKILFTSVLKEIISQTSFAIGKEKEKIVNGNLNKTDQNKELIITATDSF 179

QY 184 RMSORLITLNDTSADLMVVLPSKSLREFSAVFTDDIETVEFPSPQILFRSEHSFYTR 243
DB 180 RLSCCKIDYSN-NYNFVDVIIPKFINEIGRLISENDQVSELL-----II-----INCKQK 229

QY 244 LLEGVYDPTDRLLMTTEFEVFNVTQSLRHAMERAPLISNATQNGTVKLEITQNHISA 303
DB 230 IIEGKYPDTSKVIRTSFNSTLNINRAIKIIRRSVLSNLTMTTIVTLKIKBQKVLVTS 289

R;Fujita, M.Q.; Yoshikawa, H.; Ogasawara, N.
Mol. Gen. Genet. 215, 391-397; 1989
A;Title: Structure of the dnaA region of *Pseudomonas putida*: conservation among three
A;Reference number: Jv0002; MUID:89218947; PMID:2540413
A;Accession: Jv0002
A;Molecule type: DNA
A;Residues: 1-367 <FUJ>

RESULT 17

Nature 406, 477-483, 2000
A>Title: DNA Sequence of the cholera pathogen *Vibrio cholerae*.

A;Map position: 1

	Query Match	18.1%;	Score 336.5;	DB 2;	Length 391;	
	Best Local Similarity	26.0%;	Pred. NO. 1.1e-14;			
	Matches	101;	Conservative 79;	Mismatches 168;	Indels 41;	Gaps 12;
Qy	15	LNTTKRAISTGNAIPILSSIKIVTS--TCGVLTGSGNQGISIENTIPVSNENAGLLITSP	72			
Dd	:	:::::	:::::	:::::	:::::	:
14	LSLVSRAVSRSRPHYPVLGNVLLEADAKNYLRLTAFDLSLGIOSSFTAD-----VQGS	66				
Qy	73	GATLLRASFPFNIISLPDISINV-----KETEQHQVVLTSCKSEITLKGRDVDOYPRLL	126			
Dd	:	:::::	:::::	:::::	:::::	:
67	GRITLPAKLLNDIVSRLLPGDDITLATDPGDAGDSHLTTTISESGRFQIRGLDADFPAAL	126				
Qy	127	OEVSTENPLILKTKLKSIIAETFAASLOESRPILTGVHIVLSNKHDFKAVATDSHRWS	186			

```
Query Match      18.0%; Score 335.5; DB 2; Length 399;
Best Local Similarity 25.9%; Pred. No. 1.3e-14;
Matches 104; Conservative 82; Mismatches 171; Indels 45; Gaps 11;

QY      2 IQPSINRTLPFIHALNTTKRAISNATPILSSIKIEVITSTGVLTGNSGOISIENTIPVS 61
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      13 LKFCCLARESPASVSWAKYLPTRPTEVLGSGVLLTGSGLTISGGDFEVSAAE--VQVA 70
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      62 NENAGLLITSPGAILLLEASPFINISSLPDISINVKETEOHQWLTSGSKSEITLKGKVD 121
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      71 AE-----IASSGSVLVSGRLLSDITRALPNKPVPHP-VYDGNRVALTCCSARSLPTMAVE 124
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      122 QYPRLOEVSTENPLIIKTULLKSIIATAFPAASIQESRPILTGVHIVLSHNKDFKAVATD 181
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      125 DYPTLTPLDETG-TLPDSVFABAIGQVATAAGRDYTLPMLTIRIEISGDTWWLA-AAD 182
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      182 SHRWMSORLIILDNTSADL--MVLVPSKSLREFSAVFTDDDIETV-----EVFFSF 228
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      183 RFLAVRELKWSVLSSDFEASLVPAKTLVEVAKAGTDGSGVCULSGAGVGKQDLFGI 242
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      229 SQLFRSEHISFYTRLLEGNPYPOTDRLLIMTFEFETEVPNTQSLRHAMERAFILSNATQNG 288
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      243 SGGGKEST-----TRLLDAEFFPKRQLLPABHTAVATIDVAELTEAIKLVALV--ADRGA 295
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      289 TVKLETQNHI SAHVNSPEVKVNEDLDIVSQSGDLTISNPTVLIISLKAIKSETVKI 348
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      296 QVMREFGDGILRLSAGADDVGRABEEL-AVAFTGEPLTIAFNPNVLTDLGLASHVHSERSVF 354
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      349 HFLSPVRPPFTLTPCDEEE-----SFIQLITPVR 376
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      355 GFPTPSKPALLRPTNDVDHPTHDPGPPPALPTDYVLLMPVR 396
       ::::| | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

127 OEVTENPLILKTKLLKSIIAETAFAASLOESRPILTGVI VLSNHKDFKAVATDSHRMS 186

AB0958
DNA polymerase III beta-chain [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0958
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0958
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-366 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03158.1; PID:g16504793; GSPDB:GN00176
C:Genetics:
A:Gene: STV3941
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.8%; Score 332; DB 2; Length 366;
Best Local Similarity 24.0%; Pred. No. 2e-14;
Matches 90; Conservative 90; Mismatches 185; Indels 10; Gaps 5;

Qy 2 IQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGOISIENTIPVS 61
Db 1 MKFTVERHLKPKIQVSGPLGGRTPLFGLNLLQVADGLTSLGTDLMEWVARVTL 60

Qy 62 NENAGLLITSPGAILLEASFFINISSLIPDISINVKEIEHQHVLTSGSEITLKGD 121
Db 61 QPH-----EFGATTVPARKFDCRGLPEGAETAIVQLGDRMLVRSGRSFSLTLPAA 114

Qy 122 QYPRLOEVSTENPLILKTKLSIAETAFAASQESRPILTGV--HIVLSNHK-DFK 181
Db 115 DFNLDWQSEVEFTLPQATMKRLIEATQFSMAHQDVRYILNGM-LFETEGSELTATD 173

Qy 182 SHRMSQRLITIDNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISFY 241
Db 174 GRLIACVSMPLASLPSSHVSIVPRKGVIELMRMLDGGENPLRVQIGSNIRAHVGDFT 233

Qy 242 TRLEGNYPTDRLMTTEFEVFNTPQSLRHAMERAFILSNATQNGTVKLEITQNHISA 301
Db 234 SKLVDRGFRPDRVRLPKNPDKHLEAGCDILQAAPARAAILSNKFRG-VRLVSENQLKI 292

Qy 302 HVNSPEVKVNEEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSPVPTLTP 361
Db 293 TANNPEQEAEEILD-VSYGTEMEIGFNVSYVLDVNLKCTVVRIMLTDSVSVSQIED 351

Qy 362 GDEESFTQLITPVR 376
Db 352 A-ASQSAAYVMPMR 365

RESULT 23
AE2613
DNA polymerase III, beta chain [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AE2613
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, K.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <KUR>

A:Cross-references: UNIPROT:Q8UIJ4; GB:AB008688; PIDN:AAL41323.1; PID:g17738634; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: dnaN
A:Map position: circular chromosome
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.4%; Score 325; DB 2; Length 372;
Best Local Similarity 25.5%; Pred. No. 5.9e-14;
Matches 98; Conservative 86; Mismatches 178; Indels 22; Gaps 11;

Qy 2 IQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGOISIENTIPVS 61
Db 1 MRITLERSNLLKSLNHVHVVRERTIPILSNVLLRSGANLDMKATDLDEITEATPAM 60

Qy 62 NENAGLLITSPGAILLEASFFINISSLIPDISINVKEIEHQHVLTSGSEITLKGD 119
Db 61 VEQAG-ATTVPAAHLYE-----IVRKLDPGSEVILLATNPDGSSMTVASGRKFSLQCLP 113

Qy 120 VDQYPRLOEVSTENPLILKTKLSIAETAFAASQESRPILTGV--HIVLSNHK-DFK 176
Db 114 EADFPDLTAGTFSHTFKLKAADLQWLDRTQFAISTETRYVYLANGIFPHTTIESNGELKLR 173

Qy 177 AVATDSHRMQRLLTDLNTSADLM-VVLPKSLREFSAVFTDDIETVEVFPSPQILPVS 235
Db 174 AVATDGHRLARADVDPGSGEGMGIIPRKTGVLQKMDNPELVTVEVSDAKIRLAI 233

Qy 236 EHISFYTRLEGNYPDTDRLLMTTEFEVFNTPQSLRHAMERAFILSNATQNGTVKLEIT 295
Db 234 GSVVLTSLKLDGTQYQYRVIPTGNDKEMRVDCQTFARADVRTIS-SERGRAVKALT 292

Qy 296 QNHISAHVNSPEVKVNEEDDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSPVR 355
Db 293 DGQLTLVWNPDSGSAATEEV-AGVYDNDMSMEIGNAKYLLDITLSQSGEDA-IFLLADAG 350

Qy 356 PFTL---TPGDEESFTQLITPVR 376
Db 351 SPTLVDRDTAGDGA---LYVLMPMR 371

RESULT 24
C97395
DNA polymerase III, beta chain [imported] - Agrobacterium tumefaciens (strain C58, Cerezo)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: C97395
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens C58
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C97395
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-403 <KUR>
A:Cross-references: UNIPROT:Q8UIJ4; GB:AE007869; PIDN:AAK86116.1; PID:g15155199; GSPDB:G
C:Genetics:
A:Gene: AGR_C_520
A:Map position: circular chromosome
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.4%; Score 325; DB 2; Length 403;
Best Local Similarity 25.5%; Pred. No. 6.6e-14;
Matches 98; Conservative 86; Mismatches 178; Indels 22; Gaps 11;

Qy 2 IQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNGOISIENTIPVS 61
Db 32 MRITLERSNLLKSLNHVHVVRERTIPILSNVLLRSGANLDMKATDLDEITEATPAM 91

Qy 62 NENAGLLITSPGAILLEASFFINISSLIPDISINVKEIEHQHVLTSGSEITLKGD 119
Db 92 VEQAG-ATTVPAAHLYE-----IVRKLDPGSEVILLATNPDGSSMTVASGRKFSLQCLP 144

QY 120 VDOYPRLOEVSTENPLILKTKLLKSLIAETAFASLOESRPILTG--HIVLSNHK-DFK 176
 Db 145 EADFDELDTAGTSHTFKLKAADLKMILDTQFAITTEETRYYLNGIFFHTIESNGELKLR 204
 QY 177 AVATDSHRMSQRLITLIDNTSADLM-VVLPSKSLREFSAVFTDDIETVEFFSPSOILFRS 235
 Db 205 AVATDGHRLARADVDPAGSGECMGIIIPRKTVGELQKMDNPELEVIVESDAKIRLAI 264
 QY 236 EHISFYTRLLGNYPTDRLMTETETEVVFNQSLRHAMERAFILSNATQNGTVKLEITQ 295
 Db 265 GSVVLTKSLIDGTFFDYQVPIPTGNDKEMRVDCQTFARAVDRVSTIS-SERGRAVKLAIT 323
 QY 296 QNHISAHVNSPVGKVNEDLDIVSQSGDLTISFNPTVILIESLKAISKSETVKIHELSPVR 355
 Db 324 DQGLTIVNPNDSGATEEV-AGVYDNDMSMEIGFNAKYLDDITSQLSGEDA-IFLLADAG 381
 QY 356 PFTL---TPGDEESFIQIITPVR 376
 Db 382 SPTLVDRDTAGDDA---LYVLMFMR 402

RESULT 25
 DNA polymerase III, beta chain (dnan) RP419 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
 C:Accession: B71700
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: B71700
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-381 <AND>
 A:Cross-references: UNIPROT:Q92DB3; GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA1487
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: dnan; RP419
 C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.3%; Score 323; DB 2; Length 381;
 Best Local Similarity 24.6%; Pred. No. 8.2e-14;
 Matches 96; Conservative 87; Mismatches 182; Indels 26; Gaps 10;

QY 1 MIQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIPV 60
 Db 1 MLKLVETKTLQSLGFAISIEKENVPEYANIKLSAQDGNLESLSTNMDLYLSQKIAV 60
 QY 61 SNENAGLITSPGAILLEASFFINISSLPDISINVKEIQHVLTGSKSITILKGDV 120
 Db 61 Q-----VLNGETVATQTLSDIVRFPDPDSBELTLTEITQLBKQNCNCKFNLTLPV 113
 QY 121 DOYPRLOEVSTENPLILKTKLLKSLIAETAFASLOESRPILTG--HIVLSNHKDFKAVAT 180
 Db 114 SFPFAMDSIKPVSFKISCADPAKIESTKFSISLDETNYNLNGIYLIKD-KEPFAAST 172
 QY 181 DSHRMSQRLITLIDNTSADLMVLPKSLREFSAVFTDD---IETVEFFSPSOILF-RSE 236
 Db 173 DGYRLISWITLLEKIKFVILPQKSAEILKIVKDPKNHIEDIILLSSNKIKFICNE 232
 QY 237 HISFYTRLLGNYPTDRLMTETETEVVFNQSLRHAMERAFILSNATQNGTVKLEITQ 296
 Db 233 NTILSKLIDGTFFDYAPFAPKSSSVSKLVINRKIFADSIERTAIT-VEKPRAVKLSLR 291
 QY 297 NHI-----SAHVNSPVGKVNEDLDIVSQSGD---LTISFNPTVILIESLKAISKSETVKI 348
 Db 292 KILEISAVGEAGTAKIEITASQDKESFYENHNDESIVGFENPOYLEDVLKAIKSDIVEL 351
 QY 349 HP--LSPVRPFTLT-PGDEESFIQIITPVR 376
 Db 352 YFSDISASAPVLKIFRPNKDFIV--IMPVK 380

RESULT 26
 E72400
 DNA polymerase III, beta subunit - Thermotoga maritima (strain MS98)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: E72400
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D., C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: E72400
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-366 <ARN>
 A:Cross-references: UNIPROT:Q9WYA0; GB:AE001708; GB:AE000512; NID:g4980740; PIDN:AAD3535
 A:Experimental source: strain MS98
 C:Genetics:
 A:Gene: TM0262
 C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.1%; Score 319.5; DB 2; Length 366;
 Best Local Similarity 23.4%; Pred. No. 1.3e-13;
 Matches 85; Conservative 97; Mismatches 167; Indels 15; Gaps 9;

QY 15 LNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIPVSNENAGLITSPGA 74
 Db 14 ITIASKALAKSKVKPILAGFLFEVKGDNFYICATDLETGVRKATVNAEISGEARFVWPGD 73
 QY 75 ILLEASFFINISSLPDISINVKEIQHVLTGSKSITILKGDVQVPRLOEVSTENP 134
 Db 74 VIQK-----MVKVLPD-EITELSLEGDALVISSGSTVFRITTMPADEPFEITPAESGIT 126
 QY 135 LILKTKLKSIIAETAFASLOESRPILTG--HIVLSNHKDFKAVATDSHRMSQRLITLD 193
 Db 127 FEVDTSLLLEEMVEKVIFAAAKDEFMRNLNGVFWEL--HKNLRLVASDGFRLALAEQIE 184
 QY 194 NTSADLMVLPKSLREFSAVFTDDIE-TVEFFSPSOILFRSEHISFYTRLLGNYPT 252
 Db 185 N-EESAFLLSLKSKVEQVNDNTTEPTITVRYDGRVSLSTNDVETVMRVVDAEPDY 243
 QY 253 DRLMTETETEVVFNQSLRHAMERAFILSNATQNGTVKLEITQNHISAHVNSPVGKYN 312
 Db 244 KRVIPETETKTVVSRKELRESLKRVMVVIASKSE-SVKFEIEENVMRLVSKSPDYGEVV 302
 QY 313 EDLDIVSQSGDLTISFNPTVILIESLKAISKSETVKIHELSPVRPFTLTGDEESFIQI 372
 Db 303 DEVE-VQKEGDELVIAPNPKFIEDVLKHETEIEEMNFVDSTSPCOINPLD-ISGYLYIV 360
 QY 373 TPVR 376
 Db 361 MPVR 364

RESULT 27
 DUECB
 DNA-directed DNA polymerase (EC 2.7.7.7) III beta chain - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 28-Aug-1985 #sequence_revision 31-Mar-1998 #text_change 09-Jul-2004
 C:Accession: A91510; A93996; A22168; I41190; I54000; F65172; A00719; B24944
 R:Ohmori, H.; Kimura, M.; Nagata, T.; Sakakibara, Y.
 Gene 28, 159-170, 1984
 A:Title: Structural analysis of the dnaA and dnaN genes of Escherichia coli.
 A:Reference number: A91510; MUID:84237568; PMID:6234204
 A:Accession: A91510
 A:Molecule type: DNA
 A:Residues: 1-366 <OHM>
 A:Cross-references: UNIPROT:P00583; GB:J01602; NID:g145758; PIDN:AAB59150.1; PID:g145761
 A:Experimental source: strain K-12
 R:Blancat, M.A.; Sandler, S.J.; Armengod, M.E.; Ream, L.W.; Clark, A.J.
 Proc. Natl. Acad. Sci. U.S.A. 81, 4622-4626, 1984

A;Title: Molecular analysis of the recF gene of Escherichia coli.

A;Reference number: A93996; MUID:84272685; PMID:6379647

A;Accession: A93996

A;Molecule type: DNA

A;Residues: 297-366 <BLA>

A;Cross-references: GB:K02179; NID:g147537; PIDN:AAA24510.1; PID:g147538

R;Adachi, T.; Mizuuchi, K.; Menzel, R.; Gellert, M.

Nucleic Acids Res. 12, 6389-6395, 1984

A;Title: DNA sequence and transcription of the region upstream of the E. coli gyrB gene.

A;Reference number: A22168; MUID:84297235; PMID:6089112

A;Accession: A22168

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 342-366 <ADA>

A;Cross-references: GB:X04341; GB:X00870; NID:g41643; PIDN:CAA27869.1; PID:g41644

A;Experimental source: strain K-12

R;Armstrong, M. 263, 12109-12114, 1988

J. Biol. Chem. 263, 12109-12114, 1988

A;Title: Transcriptional Organization of the dnaN and recF Genes of Escherichia coli K-12

A;Reference number: I41190; MUID:86298898; PMID:2841344

A;Accession: I41190

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-14 <RES>

A;Cross-references: GB:M19876; NID:g145779; PIDN:AAA23695.1; PID:g551800

R;Armstrong, M.

Gene 43, 183-196, 1986

A;Title: Overlapping arrangement of the recF and dnaN operons of Escherichia coli; Position

A;Reference number: I54000; MUID:86301872; PMID:3527871

A;Accession: I54000

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 143-170 <RE2>

A;Cross-references: GB:M13822; NID:g147541; PIDN:AAA24512.1; PID:g147542

R;Blattner, F. R.; Plunkett III, G.; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; Co

.A.; Rose, D. J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: F65172

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-366 <BLAT>

A;Cross-references: GB:AR000447; GB:U00096; NID:g2367266; PIDN:AAC76724.1; PID:g1790136;

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: dnaN

A;Map position: 83 min

C;Complex: DNA polymerase III is a multichain complex; alpha, epsilon, theta, tau, gamma

A;Function: <GEN>

A;Description: DNA synthesis; synthesizes both, the lagging and the leading strands in E

A;Pathway: DNA biosynthesis

A;Note: Core enzyme (catalytic core) contains chains alpha, epsilon and theta; alpha cha

required for assembly; tau allows dimerization of the core complex and processivity is

te; the holoenzyme is completed by addition of beta chain which clamps the enzyme to DNA

C;Function: <BET>

A;Description: beta chain is required for initiation of replication

A;Note: can slide along duplex DNA bidirectionally and ATP-independent; binds core; cryst

C;Superfamily: DNA-directed DNA polymerase III beta chain

C;Keywords: DNA replication initiation; nucleotidyltransferase

Query Match

Best Local Similarity 17.1%; Score 318; DB 1; Length 366;

Matches 86; Conservative 93; Mismatches 186; Indels 10; Gaps 5;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIPVS 61

DB 1 MKFTVERHLLKPLQOVSGPLGGRTPLILGNLLQVADGTLSTGTDLMEVMVARVALV 60

QY 62 NENAGLLITSPGAILLEASFFINIITSSLPDISINVKIEHQHVLTSGKSEITLKGKDVD 121

DB 61 QPH-----EFGATTVPARKFDDICRGLPEGAIAVQLEGERMLVRSGRSRLSTLPAA 114

QY 122 QYPLQEVSTENPILITKTKLKSIIAETAPAAISOESRPILTGVHIVLSNHKPKAVATD 181

DB 115 DFPNLDDQWSEVETPLPQATMKRLIEATQFSMAHQDVRYTLNGM-LFETEGEELRTVATD 173

QY 182 SHRMSQRLITLNDTSADLMVVLPSKSLREFSAVFTDDIETVEVFFSPSQILFRSEHSIFY 241

DB 174 GHRLAVCSMPICGQSLPSHSVIVPRKGVIELMRMLDGGDNPLRVQIGSNRIAHVGDFTFT 233

QY 242 TRLEGNYPTDRLMLTMEFETEVVFNQSLRHHAMERAFNISNATQNGTVKLEITQNHISA 301

DB 234 SKLVDRGFPDYRVLPRKPKHLEAGCDLLKQAFARAAILSNEKFRG-VRLVSVSENLKI 292

QY 302 HVNSPEVKVKNEDLDIVSQGSDLTISFNPTYLIESLKAKSETVKIHLFSPVPFTLTP 361

DB 293 TANNPEQEAEEILD-VITYSGAEMEIGFNVSIVLDVNLKCNVRMMLTDSVSSVQIED 351

QY 362 GDEEESFIQLITPVR 376

DB 352 A-ASQSAAYVVMR 365

RESULT 28

D91208

DNA polymerase III beta-subunit [imported] - Escherichia coli (strain O157:H7, substrain

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: D91208

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: D91208

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-366 <HAY>

A;Cross-references: UNIPROT:P00583; GB:BA000007; PIDN:BAB38059.1; PID:g13364111; GSPDB:52

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs4636

C;Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 17.1%; Score 318; DB 2; Length 366;

Best Local Similarity 22.9%; Pred. No. 1.6e-13;

Matches 86; Conservative 93; Mismatches 186; Indels 10; Gaps 5;

QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNQISIENTIPVS 61

DB 1 MKFTVERHLLKPLQOVSGPLGGRTPLILGNLLQVADGTLSTGTDLMEVMVARVALV 60

QY 62 NENAGLLITSPGAILLEASFFINIITSSLPDISINVKIEHQHVLTSGKSEITLKGKDVD 121

DB 61 QPH-----EFGATTVPARKFDDICRGLPEGAIAVQLEGERMLVRSGRSRLSTLPAA 114

QY 122 QYPLQEVSTENPILITKTKLKSIIAETAPAAISOESRPILTGVHIVLSNHKPKAVATD 181

DB 115 DFPNLDDQWSEVETPLPQATMKRLIEATQFSMAHQDVRYTLNGM-LFETEGEELRTVATD 173

QY 182 SHRMSQRLITLNDTSADLMVVLPSKSLREFSAVFTDDIETVEVFFSPSQILFRSEHSIFY 241

DB 174 GHRLAVCSMPICGQSLPSHSVIVPRKGVIELMRMLDGGDNPLRVQIGSNRIAHVGDFTFT 233

QY 242 TRLEGNYPTDRLMLTMEFETEVVFNQSLRHHAMERAFNISNATQNGTVKLEITQNHISA 301

DB 234 SKLVDRGFPDYRVLPRKPKHLEAGCDLLKQAFARAAILSNEKFRG-VRLVSVSENLKI 292

QY 302 HVNSPEVKVKNEDLDIVSQGSDLTISFNPTYLIESLKAKSETVKIHLFSPVPFTLTP 361

DB 293 TANNPEQEAEEILD-VITYSGAEMEIGFNVSIVLDVNLKCNVRMMLTDSVSSVQIED 351

QY 362 GDEEESFIQLITPVR 376

DB 352 A-ASQSAAYVVMR 365

Db 111 PDEFPLPK--GENPEVNSAKLQAEGTLCTNGDETCLVLTGVNFKIDTNK-WQA 167
QY 178 VATDSHRMSORLITLNTSAD-LMVLPKSLRERSAVFTDDIET--VEVFFSPQILPR 234
Db 168 ASINGHKLALVCTLSEVSDPFDVTPKSLSELSKILSQSADTVSCVNILLNKTIEFS 227
QY 235 SEHISFYRLLEGNYPDTRLLMTTEFEVETVNTQSLRHAMERAFILSNATNGTVKL-- 292
Db 228 LPHTKVISRLLEGEYPKINSILPRTPEYFTLKERGFESALKRVSVLAERKQK-VVKILW 286
QY 293 EITQNHISAHVNSPEVKVNE--LDIVSQSGDLTISNPYTLIESLKAIKSETVKIHF 350
Db 287 ELEATOATLYTEATDGDVADSVLMKPAIHNSENISIGLNIDYLLGLEKHXHISTEIVVRC 346
QY 351 LSPVRPFTLTPGDEESFIQLITPV 375
Db 347 NKPTQPVICPMGGLLNQILVNPV 371

RESULT 32
AD2057
DNA polymerase III beta chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AD2057
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saamamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2057
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <KUR>
A:Cross-references: UNIPROT:Q8YVCG; GB:BA000019; PIDN:BA073709.1; PID:gl7131100; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: dnaN
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 16.9%; Score 314; DB 2; Length 387;
Best Local Similarity 25.1%; Pred. No. 3.3e-13;
Matches 92; Conservative 79; Mismatches 166; Indels 30; Gaps 10;
QY 15 LMTTKRAISTKNAIPILSSI--KIEVTGTGVTLTGSGQISIENTIPVSNENAGLLITSP 72
Db 14 LSLVNRAVPSRTHPVLNVLLQDAETNQVSLTAFDLSIGRTSF-----NAD--VWQS 66
QY 73 GAILEASFFNITISLDPISINVKI-----EQHVLTSGKSEITLKGKVDQY 123
Db 67 GAIALPAKLLVDITSRLEPEGETLDDESATDGTATGEGLI VSLTPKTKYQLRAMGAGEF 126
QY 124 PRLOQVSTENPLILKTLKLSIAETAFAASIQSRPILTGTVHIVLSNHRKDFKAVATDSH 183
Db 127 PELPLENTTAYLTATSLIEGLRSLPATGDETKQVLTGVHLTV-KQDTLEFAATDGH 185
QY 184 RMSQRLIT---LDNTSADLMVLPKSLRERSAVFTDDI---ETVEVFFSPQILPRSE 236
Db 186 RLAVVETTNERPLEDNDQOEVTPARALRELERMLAHNAASEPIALYDQGVFAWQ 245
QY 237 HISFYRLLEGNYPDTRLLMTTEFEVETVNTQSLRHAMERAFILSNATNGTVKLHI-- 294
Db 246 NORLTSRTLEGQYFAYRQLRIQPERQVTVTERRQFLSTLERIAVLAD-QKNINIVKLIDS 304
QY 295 TQNHISAHVNSPEVKVNE--LDIVSQSGDLTISNPYTLIESLKAIKSETVKIHFSPV 354
Db 305 TAQELTSLCEAQWGSGRSNV-REIAGEDIEIAFNKYLMEGLKALPSPFIQMHINQNL 363
QY 355 RPFTLTP 361
Db 364 TPVIFTTP 370

RESULT 33

AH3494

DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AH3494

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688

A:Accession: AH3494

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <KUR>

A:Cross-references: UNIPROT:Q8YED6; UNIPROT:Q8G3E6; GB:AE008917; PIDN:AAL53123.1; PID:gl

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMRI1942

A:Map position: 1

C:Superfamily: DNA-directed DNA polymerase III beta chain

C:Keywords: nucleotidyltransferase

Query Match 16.9%; Score 314; DB 2; Length 397;

Best Local Similarity 24.1%; Pred. No. 3.4e-13;

Matches 93; Conservative 89; Mismatches 178; Indels 26; Gaps 11;

QY 2 IQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTGTGVTLTGSGQISIENTIPVS 61

Db 26 MRVTLERSNLLKSLNHVHVRRERTPIILSNVLQAGASLAKMKTDLDEL----- 78

QY 62 NENAGLLITSPGAILLEASFFINIISLPD-----ISINVKIEIQHVLTTSKSEITLK 116

Db 79 NEATAAMVEOAGATTVPALHLYDIVRKLPGAEVMLSTNP---DGGSMVISGKSPRLQ 135

QY 117 GKVDVQYPRLOQVSTENPLILKTLKLSIIAETAFAASLOESRPILTV--HIVLSNHK- 173

Db 136 CLPQSDPELPTAGATPHTSFRIEQAQALKLLIDRTQFALSTETRYLNGIIFHAIESCAL 195

QY 174 DFKAATDSHRMSQ-RLITLDNTSADLMVLPKSLRERSAV--FTDDIETVEVFFSPQ 230

Db 196 KLRVATDGHRLARAELEAPSGTEGMPGIIIPKTVAELOKLVDPVDPVVTVL--SDAK 253

QY 231 ILRSEHISFYRLLEGNYPDTRLLMTTEFEVETVNTQSLRHAMERAFILSNATNGTV 290

Db 254 IRTVGSVLTSLKIDGTFPDYQYRIPSGNDKKLTIDRDFAASVDRVSTIS--SERGRAV 312

QY 291 KLEITQNHISAHVNSPEVKVNE--LDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHF 350

Db 313 KLSIADGQLTLTVNPPDGSATDEL-AADYDGPDLIGFNSKYLDD-ITGQLSGTDAVPM 370

QY 351 LSPVRPFTLTPGDEESFIQLITPV 376

Db 371 LADAGSPTLVRTDGTGDEVDLYVLMFMR 396

RESULT 34

JQ0734

DNA-directed DNA polymerase (EC 2.7.7.7) III beta chain - Proteus mirabilis

C:Species: Proteus mirabilis

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C:Accession: JQ0734

R:Skovgaard, O.

Gene 93, 27-34, 1990

A:Title: Nucleotide sequence of a Proteus mirabilis DNA fragment homologous to the 60K-r

A:Reference number: JQ0729; MUID:91033012; PMID:2172087

A:Accession: JQ0734

A:Molecule type: DNA

A:Residues: 1-367 <SKO>

A:Cross-references: UNIPROT:P22838; GB:M58352; GB:M31295; NID:gl50873; PIDN:AAA03959.1;

A:Experimental source: strain LM1509

```

Db      173 GKRLAKIDTNISLDFPSGSD--YIIPKAVEEIRMASEDVQST-IFLDQTKIAVECGNT 229
Qy      239 SFYTRLLEGNYPDRLRLMTTEFEVYVFNVTOSLRHAMERAFILSNATONGTVKLEITQNH 298
Db      230 LLVTKLISGFFPDSPVISTQSSVOLNLHREELISLLKQVALFTNESSH-SVKFQSFSPGE 288
Qy      299 ISAHVNSPEVKGKVNEDIDIVSQSGSDLTISNPTVLIESLKAIKETVKIHLFSPVRPPT 358
Db      289 LTLTANCTKVGEGKVS-M-ANNYTGETLEIAENPFFFDLKHSDRELVLQGLSDSYNPGI 347
Qy      359 LTPGDEEESFIQLITPVR 376
Db      348 IT--DSTRSLF-VIMPMR 362

RESULT 36
B41870
DNA-directed DNA polymerase (EC 2.7.7.7) III beta chain - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: B41870; T10965
R:Calcutt, M.J.; Schmidt, F.J.
J. Bacteriol. 174, 3220-3226, 1992
A:Title: Conserved gene arrangement in the origin region of the Streptomyces coelicolor
A:Reference number: A41870; MUID:92250416; PMID:1577691
A:Accession: B41870
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <CAL>
A:Cross-references: UNIPROT:P27903; GB:M82836; NID:G6539740; PIDN:AAA26735.1; J. Bacteriol. 174, 3220-3226, 1992
R:Calcutt, M.J.
Gene 151, 23-28, 1994

```

```

A/Reference number: Z17435; WOULD:95123865; PWLD:7828880
A/Accession: T10965
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 356-376 <CA2>
A/Cross-references: EMBL:L27063; NID:g436023; PID:g777777
A/Experimental source: strain A3 (2)
C/Genetics:
A/Gene: dnaN
A/Start codon: GTG
C/Superfamily: DNA-directed DNA polymerase III beta chain
C/Keywords: nucleotidyltransferase

Query Match      15.9%; Score 296; DB 2; Length 376;
Best Local Similarity 24.6%; Pred. No. 4.7e-12;
Matches 96; Conservative 77; Mismatches 184; Indels 34; Gaps 11
OV      2 I0FSINRTLFIAHNTTKRAISTKNAIPILSSIKIEVTSGTVLTGNSGOISI-----EN 56

```

1	Db	1	MKRV	RDVLAEAVAWAARS	LPAPPPAPVLA-----GLLLKAE	BGGQSLSSSPFYD	SV	51					
57	Qy	57	TIPV	SNENAGLITSP	GAILLEASFFNIIISGL	PDISINNVKEIQHVVL	TSGKSEITLK	116					
			:	:	:	:	:						
			:	:	:	:	:						
			:	:	:	:	:						
52	Db	52	SARV	SEAE---	IEEGTVL	SGRLLADISRAL	PNRPVEI--STDG	VRATVVCSSRFTLH	107				
117	Qy	117	GKVD	QVPRIQVSTEN	PLIKTKLKSIIAET	FAASLQSRSPIL	TGHI	VLSNHNKDPK	176				
108	Db	108	TLPV	EYEPALPQM--	PEATGTVPGEV	FAVOQVAIAGR	DDTL	PLVTG	VRIBIEG--DSVT	165			
			:	:	:	:	:	:	:				
			:	:	:	:	:	:	:				
177	Qy	177	AVATD	SRMSOR--	LITLDNTS	ADMLVVLPSKSL	REFSAV	TTDDLET	VEVFPSPQ---	230			
			:	:	:	:	:	:	:				
			:	:	:	:	:	:	:				
166	Db	166	LASTD	RKFAVREF	LWKPEN	PDISAV	ALVPAKTIQD	TAKALTSG	QDVLTAL	SGSGAGEGL	225		
231	Qy	231	ILFR	SEHISFYTRL	LEGNYPTD	RLLLMT	FEFEVVFNTQ	SLRHA	MEAF	LISNATQNGTV	290		
			:	:	:	:	:	:	:	:			
			:	:	:	:	:	:	:	:			
226	Db	226	IGFEG	ARRITTRL	LEGDPKYK	TLFP	TEFNS	VAVIETAP	FVEA	VKRV	ALV--AERNTPV	283	
291	Qy	291	KLEIT	QNHISAHVNS	PEGVKN	EDLDIV	QSGSD	LTISFN	PTYLIES	KAUKSETV	KIHF	350	
284	Db	284	RLSP	EOGV	LIEAGSD	DAQAV	ERVD--AQ	LEGDDIS	IAFN	TFTLLD	GLSADSP	VAQSLF	342

K850211.36
 B85533
 DNA polymerase III (beta chain) [imported] - Chlamydophila pneumoniae (strain J138)
 C/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: B85533
 K/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T
 Nucleic Acids Res. 28, 2311-2314, 2000
 A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A/Reference number: A86491, UID:20330349; PMID:10871362
 A/Accession: B85533
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-366 <STO>

```

QY      2  IQFSNRTLFHALNTTKRAISTKNAIPILSSIKIEVTSCTGVTITSGNGQISIENTIPVS 61
Db      1  MKPVSERNELGLIKIKISVPPQNTPIPLFTHVLIETYNDELVFTATDLTVTSRCVTKAK 60
QY      62  NENAGLITSPGAILLEASPTINIISLPDISINVVKIEQHVLTSGKSBITLKGKDVD 123
Db      61  -----VYEGKAISIPSKRFQVVKELTEANLEISYSAGEMAQITSGSCFRLLSMEKE 115
QY      122 QYPRLQVSTENPLTIKTKLKSIIAETAFASIQEGRPILTGVIHVLNHHKDFKAVTD 183

```

Search completed: January 28, 2005, 17:00:21
Job time : 44 secs

```
Db 114 DFMPLDQNALRSLPAEQKLTMLQRTSFVARSSESYVLTGVLAAANGV-ATVGTGTD 172
Qy 182 SHRMQ--RLITLONTSDALMLVLPKSLRFSFAVFTDDIETVEVFFSPQILFRSEHIS 239
Db 173 GKRLAKIDAETVLDKSPSG--EYIPIKAVEEIIKWCSDGEAT-IFLDQDKIAVECDNTL 230
Qy 240 FYTRLLEGNYPTDRLLMTETETEVNTQSLRHAMERAFILSNATQNGTVKLEITQNH 299
Db 231 LITKLLSGEFPDFSPVISTESNVKLDLHREELITLLKQVALFTNESSH-SVKFSFLPGL 289
Qy 300 SAHVNSPEVGVKNEDLDIVSOGSDLTISFNPTVLTIESLKAISKSETVKIHFSLSPVRP 359
Db 290 TLTANTCTKVGEGKVM-AVNTSGELLEIAFNPFPLDLKHSDKDELVSGLGSDSYNPG 348
Qy 360 TPGDEESFIQLITPVR 376
Db 349 T--DSASGLFVIMPMR 362

RESULT 40
E71559
probable DNA pol III (beta chain) - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: E71559
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: E71559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-416 <ARN>
A:Cross-references: UNIPROT:O84078; GB:AE001282; GB:AE001273; NID:G3328466; PIDN:AAC6766
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
C:Superfamily: DNA-directed DNA polymerase III beta chain

Query Match 15.7%; Score 292.5; DB 2; Length 416;
Best Local Similarity 22.7%; Pred No. 9.2e-12;
Matches 86; Conservative 91; Mismatches 183; Indels 19; Gaps 10;

Qy 1 MIQFSINRFLTHALNTTKRAITKNAIPILSSIKIEVTSTGTLTGNSGQISIENTIPV 60
Db 50 IMKFVISRNELGNLIKQNVVQSTPIPVLTVLHIESCNDELVTATDLTVSTRCVKA 109
Qy 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEIOHQVVLTSCKSEITLKGDV 120
Db 110 K-----VYESGSVTIPSRPFQILRELTAEIAEVAHSGEMATITSGSCFRLLSMGK 162
Qy 121 DOYPRLOEVSTENPLILATKLLKTIATAFAASIQESRPILITGVHIVLSNHKDFKAVAT 180
Db 163 EDFPMLPDQNALRFTLSERLKMFQRTSFVARSSESYVLTGVLAAANGV-ATVGTGTD 221
Qy 181 DSHRMQ--RLITLD-NTSADLMVLPKSLRFSFAVFTDDIETVEVFFSPQILFRSEH 237
Db 222 DGKRLAKIDTISLDPSPSGD--YIPIKAVEEIIKWCSDGEAT-IFLDQDKIAVECDNTL 278
Qy 238 ISFYTRLLEGNYPTDRLLMTETETEVNTQSLRHAMERAFILSNATQNGTVKLEITQNH 297
Db 279 TLLVTKLLSGEFPDFSPVISTESNVKLDLHREELITLLKQVALFTNESSH-SVKFSFLPGL 337
Qy 298 HISAVNSPEVGVKNEDLDIVSOGSDLTISFNPTVLTIESLKAISKSETVKIHFSLSPVR 357
Db 338 ELTANTCTKVGEGKVM-AVNTSGELLEIAFNPFPLDLKHSDKDELVSGLGSDSYNPG 396
Qy 358 TLTGDEESFIQLITPVR 376
Db 397 IIT--DSTRSLF-VIMPMR 412
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2005, 16:59:44 ; Search time 146 Seconds
(without alignments)
935.393 Million cell updates/sec

Title: US-10-048-071-28
Perfect score: 1863
Sequence: 1 MIQSFNRTLFIHALNTTKR.....LTPGDEESFIQLTPVPTN 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1849	99.2	378	15 US-10-282-122A-74301	Sequence 74301, A
2	1527	82.0	378	15 US-10-282-122A-72018	Sequence 72018, A
3	1384	74.3	378	9 US-09-815-242-13446	Sequence 13446, A
4	1384	74.3	378	15 US-10-282-122A-73682	Sequence 73682, A
5	864.5	46.4	376	9 US-09-815-242-10907	Sequence 10907, A
6	864.5	46.4	376	15 US-10-282-122A-56984	Sequence 56984, A
7	862.5	46.3	376	15 US-10-282-122A-58007	Sequence 58007, A
8	748.5	40.2	377	15 US-10-282-122A-70833	Sequence 70833, A
9	735.5	39.5	377	15 US-10-282-122A-71908	Sequence 71908, A
10	730.5	39.2	377	9 US-09-815-242-12700	Sequence 12700, A
11	730.5	39.2	377	9 US-09-815-242-12726	Sequence 12726, A
12	730.5	39.2	377	14 US-10-282-287-10	Sequence 10, Appl
13	730.5	39.2	377	15 US-10-282-122A-44084	Sequence 44084, A

ALIGNMENTS

RESULT 1

- US-10-282-122A-74301
; Sequence 74301, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

14	730.5	39.2	377	16	US-10-449-830A-2	Sequence 2, Appli
15	725.5	38.9	376	9	US-09-815-242-5669	Sequence 5669, Ap
16	703	37.7	381	15	US-10-282-122A-60333	Sequence 60333, A
17	692	37.1	379	15	US-10-282-122A-45628	Sequence 45628, A
18	586	31.5	334	15	US-10-671-403-174	Sequence 174, App
19	586	31.5	334	15	US-10-671-419-174	Sequence 174, App
20	586	31.5	334	15	US-10-670-844-174	Sequence 174, App
21	586	31.5	334	15	US-10-671-134-174	Sequence 174, App
22	586	31.5	334	15	US-10-672-098-174	Sequence 174, App
23	586	31.5	334	15	US-10-672-638-174	Sequence 174, App
24	586	31.5	334	15	US-10-673-127-174	Sequence 174, App
25	586	31.5	334	15	US-10-670-817-174	Sequence 174, App
26	586	31.5	334	15	US-10-673-119-174	Sequence 174, App
27	586	31.5	334	17	US-10-671-207-174	Sequence 174, App
28	457.5	24.6	366	15	US-10-282-122A-52020	Sequence 52020, A
29	448	24.0	402	15	US-10-282-122A-53348	Sequence 53348, A
30	442	23.7	375	15	US-10-282-122A-52377	Sequence 52377, A
31	380.5	20.4	399	15	US-10-282-122A-61654	Sequence 61654, A
32	374	20.1	387	15	US-10-282-122A-45981	Sequence 45981, A
33	364	19.5	402	9	US-09-712-363-147	Sequence 147, App
34	364	19.5	402	15	US-10-282-122A-64326	Sequence 64326, A
35	364	19.5	402	17	US-10-476-597-143	Sequence 143, App
36	363	19.5	402	15	US-10-282-122A-62773	Sequence 62773, A
37	362.5	19.5	367	15	US-10-282-122A-68144	Sequence 68144, A
38	362.5	19.5	367	15	US-10-671-403-112	Sequence 112, App
39	362.5	19.5	367	15	US-10-671-419-112	Sequence 112, App
40	362.5	19.5	367	15	US-10-670-844-112	Sequence 112, App
41	362.5	19.5	367	15	US-10-671-134-112	Sequence 112, App
42	362.5	19.5	367	15	US-10-673-098-112	Sequence 112, App
43	362.5	19.5	367	15	US-10-672-638-112	Sequence 112, App
44	362.5	19.5	367	15	US-10-673-127-112	Sequence 112, App
45	362.5	19.5	367	16	US-10-670-817-112	Sequence 112, App

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 74301
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74301

Query Match          99.2%; Score 1849; DB 15; Length 378;
Best Local Similarity 99.2%; Pred. No. 1.9e-144;
Matches 375; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
DB 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60

QY 61 SNEGALLITSPGAILLEASFFINIISLPDISINVKEIOHQVLTSGKSEITLKGKDV 120
DB 61 SNEGALLITSPGAILLEASFFINIISLPDISINVKEIOHQVLTSGKSEITLKGKDV 120

QY 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDKFKAVAT 180
DB 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDKFKAVAT 180

QY 181 DSHRMSORLITLNDTSADLMVLPKSLREPSAVFTDDIETVEVFPSPQILFRSEHISF 240
DB 181 DSHRMSORLITLNDTSADLMVLPKSLREPSAVFTDDIETVEVFPSPQILFRSEHISF 240

QY 241 YTRLLEGNYPTDRLMLTETEFVFNQTSRLHMERAFILSNATONGVKLEITQNHIS 300
DB 241 YTRLLEGNYPTDRLMLTETEFVFNQTSRLHMERAFILSNATONGVKLEITQNHIS 300

QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSPVRPFLT 360
DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSPVRPFLT 360

QY 361 PGDEESFIQLITPVRTN 378
DB 361 PGDEESFIQLITPVRTN 378

RESULT 2
US-10-282-122A-72018
; Sequence 72018, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72018
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72018

Query Match          82.0%; Score 1527; DB 15; Length 378;
Best Local Similarity 79.9%; Pred. No. 8.5e-118;
Matches 302; Conservative 39; Mismatches 37; Indels 0; Gaps 0;

QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
DB 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60

QY 61 SNEGALLITSPGAILLEASFFINIISLPDISINVKEIOHQVLTSGKSEITLKGKDV 120
DB 61 SNEGALLITSPGAILLEASFFINIISLPDISINVKEIOHQVLTSGKSEITLKGKDV 120

QY 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDKFKAVAT 180
DB 121 DOYPRLOEVSTENPLILTKLLKSIIAETAFASLQESRPILTGVHIVLSNKHDKFKAVAT 180

QY 181 DSHRMSORLITLNDTSADLMVLPKSLREPSAVFTDDIETVEVFPSPQILFRSEHISF 240
DB 181 DSHRMSORLITLNDTSADLMVLPKSLREPSAVFTDDIETVEVFPSPQILFRSEHISF 240

QY 241 YTRLLEGNYPTDRLMLTETEFVFNQTSRLHMERAFILSNATONGVKLEITQNHIS 300
DB 241 YTRLLEGNYPTDRLMLTETEFVFNQTSRLHMERAFILSNATONGVKLEITQNHIS 300

QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSPVRPFLT 360
DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFSPVRPFLT 360

QY 361 PGDEESFIQLITPVRTN 378
DB 361 PGDEESFIQLITPVRTN 378

RESULT 3
US-09-815-242-13446
; Sequence 13446, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA.011A
```

; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13446
 ; LENGTH: 378
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-09-815-242-13446

Query Match 74.3%; Score 1384; DB 9; Length 378;
 Best Local Similarity 72.2%; Pred. No. 5.8e-106;
 Matches 273; Conservative 49; Mismatches 56; Indels 0; Gaps 0;
 QY 1 MIQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNQISIENTIPV 60
 DB 1 MHFSINKNLFQALNITKRAISSKNAIPILSTVKIDVTNEGVTLTSGNQISIENTIPV 60
 QY 61 SNENAGLLITSPGAILLEASFFINISSLPDISINVKEIQHGVLTSGKSEITLKGKDV 120
 DB 61 KNEAGLLITSLGSLLEASFFINIVSSLPDVTLPKEIQHGVLTSGKSEITLKGKDV 120
 QY 121 DOYPRLOEVSTENPLILTKLLKSIIAETAPAASTQESRPILTGVHIVLSNHKDFKAVAT 180
 DB 121 EYQPRIQISASTPILLETLLKLLKIIINETAFAASTQESRPILTGVHIVLSNHKDFKAVAT 180
 QY 181 DSHRMSQRLITLNTSADLMVVLPSKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 DB 181 DSHRLSQKLLTLEKNSDDFDVVPISRLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 QY 241 YTRLLEGNYPDTDRLLMTTEFTEVFNQSLRAMEAFISNATQNGTVKLEITQNHIS 300
 DB 241 YTRLLEGNYPDTDRLLMTTEFTEVFNQSLRAMEAFISNATQNGTVKLEITQNHIS 300
 QY 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLYLESKAIKSETVKIHLFSPVPRFTLT 360
 DB 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLYLESKAIKSETVKIHLFSPVPRFTLT 360
 QY 361 PGDEEESFQILTPVVRTN 378
 DB 361 PADTDEDFMQLITPVVRTN 378

RESULT 4
 US-10-282-122A-73682
 ; Sequence 73682, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 73682
 ; LENGTH: 378
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-10-282-122A-73682
 Query Match 74.3%; Score 1384; DB 15; Length 378;
 Best Local Similarity 72.2%; Pred. No. 5.8e-106;
 Matches 273; Conservative 49; Mismatches 56; Indels 0; Gaps 0;
 QY 1 MIQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNQISIENTIPV 60
 DB 1 MHFSINKNLFQALNITKRAISSKNAIPILSTVKIDVTNEGVTLTSGNQISIENTIPV 60
 QY 61 SNENAGLLITSPGAILLEASFFINISSLPDISINVKEIQHGVLTSGKSEITLKGKDV 120
 DB 61 KNEAGLLITSLGSLLEASFFINIVSSLPDVTLPKEIQHGVLTSGKSEITLKGKDV 120
 QY 121 DOYPRLOEVSTENPLILTKLLKSIIAETAPAASTQESRPILTGVHIVLSNHKDFKAVAT 180
 DB 121 EYQPRIQISASTPILLETLLKLLKIIINETAFAASTQESRPILTGVHIVLSNHKDFKAVAT 180
 QY 181 DSHRMSQRLITLNTSADLMVVLPSKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 DB 181 DSHRLSQKLLTLEKNSDDFDVVPISRLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 QY 241 YTRLLEGNYPDTDRLLMTTEFTEVFNQSLRAMEAFISNATQNGTVKLEITQNHIS 300
 DB 241 YTRLLEGNYPDTDRLLMTTEFTEVFNQSLRAMEAFISNATQNGTVKLEITQNHIS 300
 QY 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLYLESKAIKSETVKIHLFSPVPRFTLT 360
 DB 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLYLESKAIKSETVKIHLFSPVPRFTLT 360
 QY 361 PGDEEESFQILTPVVRTN 378
 DB 361 PADTDEDFMQLITPVVRTN 378
 RESULT 5
 US-09-815-242-10907
 ; Sequence 10907, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert

61	QVDBEIVITETGSVLPGREFVDIIKKLPCKEVLKSTNEQFOTLTITGSHSEFNLSGDL	120
Qy	VDOYPLQBVSTENPLIKTLKLSIIAETAFAASLOESRPILTGVHIVLSNKKDKFVAA	179
Db	PDQYPLLPVSRDAILQISVKVLKNIIAQTWFNVSTETRPVLTVGNWLQDN-ELICTA	179
Qy	TDSHRMSQRLITLTDNTSADLMVVLPSKSLRFSFAVFTDDIETVEVFPSPQILFRSEHIS	239
Db	TDSHRLAVRKLQLEBSESNKNVILPGKALSELNKMISDSDEDIDIIPASQVLFVYGNIN	239
Qy	FYTRLEGNYPDTRILLMTEPETEVVFTQSLRHAMERAFIISNATONGTVKLEITQNH	299
Db	FISRLLEGHPDTRILFPENYEIKLGINNGDFYHAIADRASILLAREGGNNVIKLTSGNELV	299
Qy	SAHVSPEVKGYNEDLDTVSSQSDLTISFNPTVLIESLKAIKSETVKIHFILSPVRPFTL	359
Db	ELSSITSPELTGKVEEWNANDVEGGLKLSFNKMYMDALKAINDEVEVEFFCTMKPFIL	359
Qy	TPGDEESFIQIITPVRT	377
Db	KPKD-DDSVTOLILPIRT	376

```

RESULT 9
US-10-282-122A-71908
; Sequence 71908, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; FILE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71908
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71908

```

[illegible]

```

RESULT 10
US-09-815-242-12700
; Sequence 12700, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: Seq-SEQ for Windows Version 4.0
; SEQ ID NO 12700
; LENGTH: 377
; TYPE: prt
; ORGANISM: staphylococcus aureus
US-09-815-242-12700

```


QY 120 VDQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTGTVHVLNHNKDFKAVA 179
 Db 121 PDQYPLPQVSRDDAIQLSVKLVNIAQTNFVAVSTSETRPVLTVGNWLIQEN-ELICTA 179
 QY 180 TDSHRMSORLITLNTSADLMVLPKSLRPFSAVFTDDIETVEVFFSPSQILFRSEHIS 239
 Db 180 TDSHRLAVRKLOLEDVSENKNVILPGKALAEINKIMSDNEEDIDIFFASNQVLFKVGNNV 239
 QY 240 FYTRLLEGNYPDTRLLMTFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNH 299
 Db 240 FISRLLEGHYPTDTRLLFPENVEIKLSINDGFEYHAIIDRASLLAREGGNNVILKSTGDDVV 299
 QY 300 SAHNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSTVKIHLSPVRPFTL 359
 Db 300 ELSSTPEIGTVKEEVANDVEGSLKISFNSKYMMDALKAINDEVEVEFFGTMKPFIL 359
 QY 360 TP-GDEEESFIQLITPVRT 377
 Db 360 KPKGDD--DSVTQLILPIRT 376

RESULT 13

US-10-282-122A-44084
 ; Sequence 44084, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 44084
 ; LENGTH: 377
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-282-122A-44084

Query Match 39.2%; Score 730.5; DB 15; Length 377;
 Best Local Similarity 39.3%; Pred. No. 7e-52;

Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNSGOISIENTIPV 60
 Db 1 MMEFTIKRDYFITQLNDTLKAISSPRITLITGKIDAKEHEVILTGSDSEISIEITIPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLPDISINVKIEIQHVVLTSKGSEITLKGKD 119
 Db 61 TVDGEDIVNISSETGSVLPGRFFVDIIKKLPKGVKLTSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTGTVHVLNHNKDFKAVA 179
 Db 121 PDQYPLPQVSRDDAIQLSVKLVNIAQTNFVAVSTSETRPVLTVGNWLIQEN-ELICTA 179
 QY 180 TDSHRMSORLITLNTSADLMVLPKSLRPFSAVFTDDIETVEVFFSPSQILFRSEHIS 239
 Db 180 TDSHRLAVRKLOLEDVSENKNVILPGKALAEINKIMSDNEEDIDIFFASNQVLFKVGNNV 239
 QY 240 FYTRLLEGNYPDTRLLMTFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNH 299
 Db 240 FISRLLEGHYPTDTRLLFPENVEIKLSINDGFEYHAIIDRASLLAREGGNNVILKSTGDDVV 299
 QY 300 SAHNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSTVKIHLSPVRPFTL 359
 Db 300 ELSSTPEIGTVKEEVANDVEGSLKISFNSKYMMDALKAINDEVEVEFFGTMKPFIL 359
 QY 360 TP-GDEEESFIQLITPVRT 377
 Db 360 KPKGDD--DSVTQLILPIRT 376

RESULT 14

US-10-449-830A-2
 ; Sequence 2, Application US/10449830A
 ; Publication No. US20040137516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Phagotech Inc.
 ; APPLICANT: Pelletier, Jerry
 ; APPLICANT: Dubow, Michael
 ; APPLICANT: Gros, Philippe
 ; APPLICANT: Bergeron, Dominique
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS AU
 ; TITLE OF INVENTION: AND ITS ENCODED PROTEIN STAAU_R2
 ; FILE REFERENCE: Q79017
 ; CURRENT APPLICATION NUMBER: US/10/449,830A
 ; CURRENT FILING DATE: 2003-05-31
 ; PRIOR APPLICATION NUMBER: PCT/CA01/01754
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: US 09/727,892
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: US 09/885,561
 ; PRIOR FILING DATE: 2001-06-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 377
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-449-830A-2

Query Match 39.2%; Score 730.5; DB 16; Length 377;

Best Local Similarity 39.3%; Pred. No. 7e-52;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIQFSINRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGNSGOISIENTIPV 60
 Db 1 MMEFTIKRDYFITQLNDTLKAISSPRITLITGKIDAKEHEVILTGSDSEISIEITIPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLPDISINVKIEIQHVVLTSKGSEITLKGKD 119
 Db 61 TVDGEDIVNISSETGSVLPGRFFVDIIKKLPKGVKLTSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTGTVHVLNHNKDFKAVA 179


```

Db 121 PQYPLLQVSDRDAIQSVKLVKNVIAQTNFAVSTSRPVLTVGNWLIQEN-ELICTA 179
QY 180 TDSHMSORLIITDNTSADLMVLPKSLRSEFSAVFTDDIETVEVFFSPSQILFRSEHIS 239
Db 180 TDSHLAVRKQLQEDVSEKNVLIIFCKALAEINKIMSDNEEDIDIFFASNQVLPKVGNNV 239
QY 240 FYTRLLEGNYPTDRLMTTEFTEVFNQSRHMERAFNISNATQGTVKLEITQNH 299
Db 240 FISRLLEGHYPTDRLFPENYIKLSIDNGEYHAIDRASLLAREGGNNVILKSTGDDVV 299
QY 300 SAHVNSPEVGVKNEEDLIVSQSGDLTISFNPTYLIESIKATKSETVKIHFELSPVRPTL 359
Db 300 ELSTSPETIGTVKEEVANDVEGGLKISFNKYMMDALKADNDEVEVEFGTMMKPFIL 359
QY 360 TP-GDEEESFIQLITPVRT 377
Db 360 PKPGD--DSVTQLILPIRT 376

```

RESULT 15

```

US-09-815-242-5669
; Sequence 5669, Application US/09815242
; Publication No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5669
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5669

```

```

Query Match 38.9%; Score 725.5; DB 9; Length 376;
Best Local Similarity 39.2%; Pred. No. 1.8e-51;
Matches 148; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 2 IOFSNRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTVGSGNGOISIENTIPVS 61
Db 1 MEFTIKRDYFIQLNDTLKALISPRITLPILTGKIDAKEHEVILTVGSGSEISIEITIPKT 60
QY 62 NENAGL-ITSPGAILLEASFFINISSLIPDISINVKIEQHQVVLTSKSEITLKGKDV 120
Db 61 VGDGDIVNISGTVLPGRFFVDIIKKLPKDKVLKSTNEQFQTLTSGHSEFNLSGLDP 120
QY 121 DOYPRLOEVSTENPILKTLKLSIAETAFAASIQESRPILTGVHVLVSNHOKFAVAT 180

```

```

Db 121 DOYPLLQVSDRDAIQSVKLVKNVIAQTNFAVSTSRPVLTVGNWLIQEN-ELICTAT 179
QY 181 DSHMSORLIITDNTSADLMVLPKSLRSEFSAVFTDDIETVEVFFSPSQILFRSEHISF 240
Db 180 DSHRLAVRKQLQEDVSEKNVLIIFCKALAEINKIMSDNEEDIDIFFASNQVLPKVGNNV 239
QY 241 YTRLLEGNYPTDRLMTTEFTEVFNQSRHMERAFNISNATQGTVKLEITQNHIS 300
Db 240 ISRLLEGHYPTDRLFPENYIKLSIDNGEYHAIDRASLLAREGGNNVILKSTGDDVVE 299
QY 301 AHVNSPEVGVKNEEDLIVSQSGDLTISFNPTYLIESIKATKSETVKIHFELSPVRPTL 360
Db 300 LSSTSPETIGTVKEEVANDVEGGLKISFNKYMMDALKADNDEVEVEFGTMMKPFILK 359
QY 361 P-GDEEESFIQLITPVRT 377
Db 360 PKPGD--DSVTQLILPIRT 375

```

RESULT 16

```

US-10-282-122A-60333
; Sequence 60333, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60333
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60333

```

```

Query Match 37.7%; Score 703; DB 15; Length 381;
Best Local Similarity 38.8%; Pred. No. 1.3e-49;
Matches 148; Conservative 93; Mismatches 134; Indels 6; Gaps 5;

```



```
Query Match 31.5%; Score 586; DB 15; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINISSLPDISINVKIEHQHVVL 106
Db 1 NSDIIIESFIPLEKEGKLLVDVKRPGSIVLQARFSEIVKLPQQTVEIETEDNFLTII 60

QY 107 TSGKSEITLKGKDQVDPYRLOEVSTENPILKTKLLKSIITAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRNLNGNADEYPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120

QY 167 IVLSNHKDFKAVATDSHRMSQRLITLDNTSADLM--VVLPSKLSREFSAVFTDD-IETVE 223
Db 121 WKV-EHGLVCTATDSHRLAMRKVIIIESENEVSNVIVPGKSLNELSKIIILDDGNHPVD 179

QY 224 VFFSPSQILFRSEHISFYTRLLEGNYPTDRLMLTTEFEFTEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAEHLFFSRLLDGNYPETARLIPTESKITMIVNAKEFLQAIIDRASLLAR 239

QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPDGMLEISSISPEIGKVTQLOQTESLEGEELNISFSKAYMMDALRALD 299

QY 343 SETVKIHLFSPVRPTLTPGDEEBSFIQIITPVRT 377
Db 300 GTDIIQISFTGMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 19
US-10-671-419-174
; Sequence 174, Application US/10671419
; Publication No. US20040038290A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/671,419
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-10-671-419-174

Query Match 31.5%; Score 586; DB 15; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINISSLPDISINVKIEHQHVVL 106
Db 1 NSDIIIESFIPLEKEGKLLVDVKRPGSIVLQARFSEIVKLPQQTVEIETEDNFLTII 60

QY 107 TSGKSEITLKGKDQVDPYRLOEVSTENPILKTKLLKSIITAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRNLNGNADEYPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120

QY 167 IVLSNHKDFKAVATDSHRMSQRLITLDNTSADLM--VVLPSKLSREFSAVFTDD-IETVE 223
Db 121 WKV-EHGLVCTATDSHRLAMRKVIIIESENEVSNVIVPGKSLNELSKIIILDDGNHPVD 179

QY 224 VFFSPSQILFRSEHISFYTRLLEGNYPTDRLMLTTEFEFTEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAEHLFFSRLLDGNYPETARLIPTESKITMIVNAKEFLQAIIDRASLLAR 239

QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPDGMLEISSISPEIGKVTQLOQTESLEGEELNISFSKAYMMDALRALD 299

QY 343 SETVKIHLFSPVRPTLTPGDEEBSFIQIITPVRT 377
Db 300 GTDIIQISFTGMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 19
US-10-671-419-174
; Sequence 174, Application US/10671419
; Publication No. US20040038290A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/671,419
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-10-671-419-174

Query Match 31.5%; Score 586; DB 15; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINISSLPDISINVKIEHQHVVL 106
Db 1 NSDIIIESFIPLEKEGKLLVDVKRPGSIVLQARFSEIVKLPQQTVEIETEDNFLTII 60

QY 107 TSGKSEITLKGKDQVDPYRLOEVSTENPILKTKLLKSIITAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRNLNGNADEYPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120

QY 167 IVLSNHKDFKAVATDSHRMSQRLITLDNTSADLM--VVLPSKLSREFSAVFTDD-IETVE 223
Db 121 WKV-EHGLVCTATDSHRLAMRKVIIIESENEVSNVIVPGKSLNELSKIIILDDGNHPVD 179

QY 224 VFFSPSQILFRSEHISFYTRLLEGNYPTDRLMLTTEFEFTEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAEHLFFSRLLDGNYPETARLIPTESKITMIVNAKEFLQAIIDRASLLAR 239

QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPDGMLEISSISPEIGKVTQLOQTESLEGEELNISFSKAYMMDALRALD 299

QY 343 SETVKIHLFSPVRPTLTPGDEEBSFIQIITPVRT 377
Db 300 GTDIIQISFTGMRPFLLRP-LHTDSMLQLILPVRT 333
```



```
; CURRENT APPLICATION NUMBER: US/10/672,638
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-10-673-127-174

Query Match      31.5%; Score 586; DB 15; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPLDISINVKEIQHVVL 106
DB 1 NSDISIIESFPLEKEGKLLVDVVRPGSIVLOARFSEIVKLPQOTVEIETEDNFLI 60
QY 107 TSGKSEITLKGKVDQYPRLOEVSTENPLILKTLKLSIIAETAPASLOESRPILTVGH 166
DB 61 RSGHSEFRLGNLADEXPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSQRLITLNTSADLM--VVLPSKSLREPSAVFTDD-IETVE 223
DB 121 WKV-EHGELVCTATDSHRLAMRKVKIIESENEVSYNVIPGKSLNELSKILLDDGNHPVD 179
QY 224 VFFSPQILFRSEHSFYTRLLEGNYPDTRLLMTTEFEVVFNTQSLRHAMERAFILSN 283
DB 180 IVMANQVLFKAEHLFFSRLLDGNYPETARLIPETSKITMIVNAKEFLQAI DRASLLAR 239
QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKVNEDLDIVSOGSDLTISFNPTYLIESLKA 342
DB 240 EGRNVVVKLTLPGGMLEISSISPEIGKVTEQLQTESLEGEELNIFSASKYMMDALRALD 299
QY 343 SETVKIHLSPVRPTLTPGDEESFQIOLITPVRT 377
DB 300 GTDIQISPTGAMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 24
US-10-673-127-174
; Sequence 174, Application US/10673127
; Publication No. US20040081995A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/673,127
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-10-673-127-174

Query Match      31.5%; Score 586; DB 16; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPLDISINVKEIQHVVL 106
DB 1 NSDISIIESFPLEKEGKLLVDVVRPGSIVLOARFSEIVKLPQOTVEIETEDNFLI 60
QY 107 TSGKSEITLKGKVDQYPRLOEVSTENPLILKTLKLSIIAETAPASLOESRPILTVGH 166
DB 61 RSGHSEFRLGNLADEXPRLPQIEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSQRLITLNTSADLM--VVLPSKSLREPSAVFTDD-IETVE 223
DB 121 WKV-EHGELVCTATDSHRLAMRKVKIIESENEVSYNVIPGKSLNELSKILLDDGNHPVD 179
QY 224 VFFSPQILFRSEHSFYTRLLEGNYPDTRLLMTTEFEVVFNTQSLRHAMERAFILSN 283
DB 180 IVMANQVLFKAEHLFFSRLLDGNYPETARLIPETSKITMIVNAKEFLQAI DRASLLAR 239
QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKVNEDLDIVSOGSDLTISFNPTYLIESLKA 342
DB 240 EGRNVVVKLTLPGGMLEISSISPEIGKVTEQLQTESLEGEELNIFSASKYMMDALRALD 299
QY 343 SETVKIHLSPVRPTLTPGDEESFQIOLITPVRT 377
DB 300 GTDIQISPTGAMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 24
US-10-673-127-174
; Sequence 174, Application US/10673127
; Publication No. US20040081995A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; FILE REFERENCE: 22221/1030
; CURRENT APPLICATION NUMBER: US/10/673,127
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
```

```

QY 49 NGOIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPDISINVKIEHQHVUL 106
Db 1 NSDISIIESFIPLEKEGKLLVDVVRPGSVIVLQARFSEIVKVLPOQVTEIETDNFLTII 60
QY 107 TSGKSEITLKGDVDQVPRLOEVSTENPLILKTLKLSIIAETAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRLNGLNADEYPRLPQIEEENVFOIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSORLITLNTSADLM--VVLPSKSLRSEPSAVFTDD-IETVE 223
Db 121 WKV-EHGELVCTATDSHRLAMRKVKIIESENEVSNVIVPGKSLNELSKILLDDGNHPVD 179
QY 224 VFPSQILFRSEHISFYTRLLEGNYPDTRLMLTEFETEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAHLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQADIRASLLAR 239
QY 284 ATONGTVKL-EITQNHISAHVNSPEVKYNEDLDIVSQSGDLTITSFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPFGMLLEISSISPEIGKVTEQLQTESLEGEELNIFSASKYMMDALRALD 299
QY 343 SETVKHFLSPVRPFTLTPGDEESFIQLITPVRT 377
Db 300 GTDIQISFTGAMRPFLLRP-LHTDSMLQLILPVRT 333

```

RESULT 26

```

US-10-673-119-174
; Sequence 174, Application US/10673119
; Publication No. US20040110210A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 22221/1030
; CURRENT FILING DATE: 2003-09-26
; PRIOR FILING DATE: 2003-09-26
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRF
; ORGANISM: Bacillus stearothermophilus
US-10-673-119-174

```

```

Query Match 31.5%; Score 586; DB 16; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NGOIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPDISINVKIEHQHVUL 106
Db 1 NSDISIIESFIPLEKEGKLLVDVVRPGSVIVLQARFSEIVKVLPOQVTEIETDNFLTII 60
QY 107 TSGKSEITLKGDVDQVPRLOEVSTENPLILKTLKLSIIAETAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRLNGLNADEYPRLPQIEEENVFOIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSORLITLNTSADLM--VVLPSKSLRSEPSAVFTDD-IETVE 223
Db 121 WKV-EHGELVCTATDSHRLAMRKVKIIESENEVSNVIVPGKSLNELSKILLDDGNHPVD 179

```

```

QY 224 VFPSQILFRSEHISFYTRLLEGNYPDTRLMLTEFETEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAHLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQADIRASLLAR 239
QY 284 ATONGTVKL-EITQNHISAHVNSPEVKYNEDLDIVSQSGDLTITSFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPFGMLLEISSISPEIGKVTEQLQTESLEGEELNIFSASKYMMDALRALD 299
QY 343 SETVKHFLSPVRPFTLTPGDEESFIQLITPVRT 377
Db 300 GTDIQISFTGAMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 27
US-10-671-207-174
; Sequence 174, Application US/10671207
; Publication No. US20040197796A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 22221/1030
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-11-21
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1997-04-08
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 174
; LENGTH: 334
; TYPE: PRF
; ORGANISM: Bacillus stearothermophilus
US-10-671-207-174

```

```

Query Match 31.5%; Score 586; DB 17; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.3e-40;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;

QY 49 NGOIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPDISINVKIEHQHVUL 106
Db 1 NSDISIIESFIPLEKEGKLLVDVVRPGSVIVLQARFSEIVKVLPOQVTEIETDNFLTII 60
QY 107 TSGKSEITLKGDVDQVPRLOEVSTENPLILKTLKLSIIAETAFAASLOESRPILTVGH 166
Db 61 RSGHSEFRLNGLNADEYPRLPQIEEENVFOIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSORLITLNTSADLM--VVLPSKSLRSEPSAVFTDD-IETVE 223
Db 121 WKV-EHGELVCTATDSHRLAMRKVKIIESENEVSNVIVPGKSLNELSKILLDDGNHPVD 179
QY 224 VFPSQILFRSEHISFYTRLLEGNYPDTRLMLTEFETEVVNTQSLRHAMERAFILSN 283
Db 180 IVMTANQVLFKAHLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQADIRASLLAR 239
QY 284 ATONGTVKL-EITQNHISAHVNSPEVKYNEDLDIVSQSGDLTITSFNPTYLIESLKAIK 342
Db 240 EGRNVVVKLTLPFGMLLEISSISPEIGKVTEQLQTESLEGEELNIFSASKYMMDALRALD 299
QY 343 SETVKHFLSPVRPFTLTPGDEESFIQLITPVRT 377

```

```

Db      300 GTDIOISFTGMRPFLLRP-LHTDSMLQLILPVRT 333

RESULT 28
US-10-282-122A-52020
; Sequence 52020, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 52020
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-52020

Query Match      24.6%; Score 457.5; DB 15; Length 366;
Best Local Similarity 29.9%; Pred. No. 2.6e-29;
Matches 112; Conservative 91; Mismatches 159; Indels 13; Gaps 7;

QY      2  IQFSNRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSNGQISIENTIPVS 61
Db      1  MFCEKNTLQRAITIAQAVTKSTMPVLQGLMSVQNNELTIGSDIDLSIETKINVE 60

QY      62  NENAGLITSPGAILLEASFFNIITSSLPDISINVKIEHQHQLVLTSGKSEITLKGKQVD 121
Db      61  -----VLEEGKVLDAKLLSEIRKLPNSKVEIQTIENNVCVEITCNKSKLTLVVLNPN 113

QY      122  QYRPLQEVSTENPLILKTLKLSIAETAFAASLOESRILTCGVHVLNHNKDFKAVATD 181
Db      114  DPFSLPEIDENSIFKINQKTLTKMTIKGTIFAQAQDETFRILGTGVFEIKDKS-LNLVAID 172

QY      182  SHRMSORLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFSPSOILFRSEHISFY 241
Db      173  GYRLAURSQYIDNETS-INAVIPGKTLNEVIKILEDG-D-VNITFTSNHILFNLGNTKII 230

```

```

QY      242  TRLEGNYPDTRLLMTTEFEVTVNTQSLRHAMERAPFLISNATONGTVKLEITQNHISA 301
Db      231  SRLLEGEFIKYNISIIPEEYNLNIIVARKEELDCIERASIMAKOGNNNIKLDI-EDDVMY 289

QY      302  HVNSEVGVKVNEDLDIVSQSGDLTISFNPTVLTIESKAIKSETVKIHFLSVPRFTLTP 361
Db      290  ITSNSQLGNVREINIIQ-GQPLKIAFNKYKLLIDVLKIMNQEEIVMFMSSSISCIIXN 348

QY      362  GDEESFTQLITPVR 376
Db      349  KENDDSD-TYLILPVR 362

RESULT 29
US-10-282-122A-53348
; Sequence 53348, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53348
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-282-122A-53348

Query Match      24.0%; Score 448; DB 15; Length 402;
Best Local Similarity 28.8%; Pred. No. 1.8e-28;
Matches 107; Conservative 96; Mismatches 153; Indels 16; Gaps 7;

QY      7  NRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSNGQISIENTIPVSNENAG 66
Db      40  NQKILANRIGIAQAKINGKTTIELLKILISTEEGQLKLTGYDAEIGIETVYQAE----- 94

QY      67  LLITSPGAILLEASFFNIITSSLPDISINVKIEHQHQLVLTSGKSEITLKGKQVDQYPL 126

```

Db 95 --LIEKGVVVVVDARLFGDIIRKLPDSFVEIETDSENNIYNCVNSRPFKIGYAAKEFPKL 152
 QY 127 QEVSTENPLIKTKLJKLSIIAETAFAASLQESRPILT--VHIVLSNKHDKRAVATDSHR 184
 Db 153 PELNEEDLYSIPQELIKNMIKQIVFALSQDQKFPVLMGELLEIV---DRNLNLVAIDGVR 209
 QY 185 MSORLTLNDSADLMVLPSPKSLREFSAVFTDDIETVEFFSPSQILFRSEHISFYTRL 244
 Db 210 LAVKSCSVDSLNIKVIIPGKTLIDVNSLLSGE-DNVKVGFEKNAIFIINDTKIITRL 268
 QY 245 LEGNYPDTRLMLTEFETEVNTQSLRHAMERAFILSNATQNGTVKLEITQNHISAHVN 304
 Db 269 LEGDFIDYKCLLPRENHRSVKLTKELLNRSASLLSQSEKNLKLKLSI-RDKWMAITS 327
 QY 305 SPEVGVKVNEDLIVSQSGDLTISFNPTYLIESLKAIKSETVKIHLFSLVPRFPTLPDGE 364
 Db 328 NTEKGNVYEVEI-DLDGDLDFIAFNSRYFIEGLKNIDNEEFIEFTTNVNCIIPKPTD- 385
 QY 365 ERSFQIITPVR 376
 Db 386 DVNIYIYLLPVR 397

RESULT 30

US-10-282-122A-52377
 ; Sequence 52377, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA 034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 52377
 ; LENGTH: 375
 ; TYPE: PRT
 ; ORGANISM: Clostridium botulinum
 US-10-282-122A-52377

Query Match 23.7%; Score 442; DB 15; Length 375;
 Best Local Similarity 29.4%; Pred No. 5.2e-28;
 Matches 113; Conservative 79; Mismatches 165; Indels 28; Gaps 8;
 QY 2 IQPSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSNGQISIENTIPVS 61
 Db 9 MKITCQKNILLEGISIVQKAVTGKSTLPILSGILIRANKNELVLTGSDWDLSTETKVKAN 68
 QY 62 NENAGLLTSCAILLEASFNITISSLPDISINVKEIQHVLTSGKSEITLKGKVDV 121
 Db 69 -----LLEGTIVLDARLFGDIIRKLPNDLIEINTLSDNSIEIICQNSRNLNMPG 121
 QY 122 QYPRLOQEVSTENPLIKTKL-----LKSIIAETAFAASLQESRPILTGVHIVLSNKH 173
 Db 122 EFP-----NPPILNENMIFSGESKLNWKGTIFATAQDETRPILTVLGVFOIKD-K 172
 QY 174 DFKAVATDSHRMSQRLITLNTSADLMVLPSPKSLREFSAVFTDDIETVEFFSPSQILF 233
 Db 173 MLNLVALDGYRLALRSEVVDNDNT-INAVIPGKTLNEVSKILEEDNENVNITFTPNHLF 231
 QY 234 RSEHISFYTRLLEGNYPTDRLMLTEFETEVNTQSLRHAMERAFILSNATQNGTVKLE 293
 Db 232 SIGETKIISRLLEGEFISYKSIIPPEFNKLIIRAKRSELLNRSIERASLMAKEGNTLVKFPD 291
 QY 294 ITQNHISAHVNSPEVGVKVNEDLIVSQSGDLTISFNPTYLIESLKAIKSETVKIHLFSLP 353
 Db 292 FSDDKIVITNS-QLGWVREELKVVLO-GEQDQIAFNSKYLLDLVLTMEDNEVVLFEFSSS 349
 QY 354 VRPFTLPDGEDEESFQIITPVRN 378
 Db 350 VSP-CIIKNTEINNCTVLVLPVRLN 373

RESULT 31

US-10-282-122A-61654
 ; Sequence 61654, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA 034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308


```
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61654
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-61654

Query Match          20.4%; Score 380.5; DB 15; Length 399;
Best Local Similarity 26.1%; Pred. No. 7e-23;
Matches 104; Conservative 87; Mismatches 170; Indels 37; Gaps 10;

QY 2 IQFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
13 LKFLRVRESFADAVSWAKSLPSRPAPVLSGVLLSGTDEGLTISGDFYEVSAEAQVAE 72
QY 62 NENAGLLITSPAILLEASFFNIISLPPDISINVKIEIQHOVLTSGKSEITLKGKVD 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
73 -----IASPGSVLVSGRLSDIVRALPNKPIDF-YVDGNRVALLNCGSARFSLPTMAVE 124
QY 122 QYRPLQEVSTENPLILKTLKLSIIAETAFAASLOESRPILTGTVHVLNHNKDKFRAVATD 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 DYTTLPTPEETG-TLPADLPFAEATGQVAIAAGRDDTLPMLTGIRVEISGDTVVLA-ATD 182
QY 182 SHRMSORLITLNTSADL--MVLPSKSLREFSAVFTDDIETVEVFFSPQILFRSEHIS 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
183 RPLAVRELTWASAPDIEAAVLVPAKTLAAEARTGIDGSD-VRLSLGAGAGVGKDGLLG 241
QY 240 F-----YTRLEGNYPTDRLMTETETEVVFNQSLRHAMERAFILSNATQNGTVKL 292
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
242 ISGNKGRSTRLLDAEFKFRQLLPAETHAVATINVAELTEAIKLVALLV--ADRGAAQVRM 299
QY 293 BITONHISAHVNSPEVKVNEOLDIVSQSGSLTISFNPTYLIESLKAIKSETVTKIHFLS 352
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 EFSEGLSLRISAGADDVGRAEEDL-AVDFAPELTTAFNPTLYTLTDGLGSRVSERSVFGTT 358
QY 353 PVRPFTLTPGDEES-----FIQLITPVR 376
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 PGKALLRPASDDSPPSGSGFSPALPTDVTYVLLMPVR 396

RESULT 32
US-10-282-122A-45981
; Sequence 45981, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
```

```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45981
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (32)..(32)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (56)..(56)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (75)..(75)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (109)..(109)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (199)..(199)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (289)..(289)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (296)..(296)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (304)..(304)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (328)..(328)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-45981

Query Match          20.1%; Score 374; DB 15; Length 387;
Best Local Similarity 29.2%; Pred. No. 2.3e-22;
Matches 117; Conservative 72; Mismatches 170; Indels 42; Gaps 14;

QY 4 PSINRFLTHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLT-TGSGQISIENTIPVSN 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 FSIQKAYLVRSVQDVMKAVSFRTTTPILTGTIXVWASGKSPLYQGSABEISNGNFXPHLK 60
QY 63 ENAGLLI--TSPGAILLEASFFNIISLDP---ISINVKIEIQHOVLTSGKSEITLTK 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 RMGKIVKVGSGKXKWFRAKNFSENGKLPKRWKEKFEVNLFGQXKLFGEIHNFVK 120
QY 117 GKVDVQYRPLQEVSTENPLILKTLKLSIIAETAFAASLOESRPILTGTVHVLNHNKDKF 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 WLDSAEYPLLQIEBHVFKEIPTDLLKHMIRQTVFAVSTETRPILTGV-----NWKVYT 175
QY 177 A----VATDSHRMSORLITLNTSADL--MVLPSKSLREFSAVFTDDIETVEVFFSPSQ 230
```

Db 176 SGLNLYCTDSRLALRAKIBGMXDRPRKPFVIGKSNELSKLDSEEWIVITEYQ 235
QY 231 ILFSEHISFVTRLLGNYDPTDLLMTE-FETEVWNTQ-SURHAMERAFILSNATQNG 288
Db 236 VLFRTKHLFFSKIVRRKYPDTLLISCRGVKQDFLVNYQKNFYKAIDS--VISVXLENG 293
QY 289 -----TVKLEITQNHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIE 336
Db 294 RNXCCEIYRLXQWALBIC-SEFAGNMGVGGGSIXK-LD-----GEEIKISPSAKYMD 346
QY 337 SLKAIKETVKIHLPSVRFTLTPGDEESFIQLITPVRT 377
Db 347 ALKALDSTEIKISFTGAMRPFILRTVN-DESIQLILPVRT 386

RESULT 33

US-09-712-363-147
; Sequence 147, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09712,363
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-147

Query Match 19.5%; Score 364; DB 9; Length 402;
Best Local Similarity 25.8%; Pred. No. 1.7e-21;
Matches 106; Conservative 82; Mismatches 163; Indels 60; Gaps 11;
QY 2 IOFSINRTLFIHALNTTKRAISTKNAIPILSIIKIEVTSTGVTLTSGNGQISIENTIPVS 61
Db 13 LTFRLRESFADAVSWAKNLPARPAVPLVSGVLTGSDNGLTISGFDYVSAEQV--- 69
QY 62 NENAGLLTSPAILLEASFNNIISLPDISINVKEIQHVLTSGKSIITLKGKQVD 121
Db 70 ----GAEIVSPGSLVSGRLSDITRALPNKFPDV-HVEGNRVALTTCGNARFSLPTWPE 124
QY 122 QYRPLQEVSTENPLIKTLKLSIIAETAFAASLQESRPILTGVHIVLSNKHDKFAYATD 181
Db 125 DYPVLTPLPEETGL-LPAELFAEALSQVIAAGRDDTLPLMIGIRVELGTTVVLA-ATD 182
QY 182 SHRMSQRILTDNTSADL--MVVLPSKSLRBS-----AVFTDDIETV 222

Db 183 RFLAVRELKWSASSPDIEAAVLVPAKTLAABAAGICGSDVRLSLGTGPGVGKGLLGI 242
QY 223 EVFFSPSQILFRSEHISFVTRLLGNYDPTDLLMTEFETEVWNTQSLRHAMERAFILIS 282
Db 243 SCNGKRST-----TRLLDAEFFKPFQLLPTETAATVATMDVAELIEAIKLVAVL- 290
QY 283 NATQNGTVKLEITQNHISAHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIK 342
Db 291 -ADRGAVRMFADGVSRLSAGADDVGRABEDL-VVDYAGEPLTIAFNPTVLTDLGSLSLR 348
QY 343 SETVKIHLSPVRPPTLTP--GDE-----EESFIQLITPVR 376
Db 349 SERVSGFTTAGKPALLRPVSGDDPRVAGLNGNGPFPAPVSTDYVYLLMPVR 399

RESULT 34

US-10-282-122A-64326
; Sequence 64326, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64326
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64326

Query Match 19.5%; Score 364; DB 15; Length 402;
Best Local Similarity 25.8%; Pred. No. 1.7e-21;
Matches 106; Conservative 82; Mismatches 163; Indels 60; Gaps 11;
QY 2 IOFSINRTLFIHALNTTKRAISTKNAIPILSIIKIEVTSTGVTLTSGNGQISIENTIPVS 61
Db 13 LTFRLRESFADAVSWAKNLPARPAVPLVSGVLTGSDNGLTISGFDYVSAEQV--- 69

QY 62 NENAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGKQVD 121
Db 70 ----GAEIVSPGSLVSGRLSDITRALPNKQVDV-HVEGNRVALTTCGNARFSLTMPVE 124
QY 122 QYPRLOEVSTENPLIKTKLLKSIITAEATAFAASLOESRPLTGVHVLNHNKDKFVATD 181
Db 125 DYPTLTPTPEETGL-LPAELFAEASQVAIAAGRDITLPMLTGIRVILGETVWLA-ATD 182
QY 182 SHRMSQRLITLNTSADL--MVLPSKSLREFS-----AVFTDDIETV 222
Db 183 RFLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAGIGSDVRLSLGTCGPGVKDGLGI 242
QY 223 EVFFSPSQILFRSEHISFYTRLEGNYPDTDRLLMTETETEVVFNQSLRHAMERAFILIS 282
Db 243 SGNKRST-----TRLLDAEFKFRQLLPTEHTAVATMDVABIEAIAKLVALV- 290
QY 283 NATONGTVKLEITQNHISAHVNSPEVGKVNEDLDIVSOGSDLTISFNPTYLIESLKAIK 342
Db 291 -ADGAQVRMEFADGSRVLSAGADVDGRAEBDL-VVDYAGEPLTIAFNPTYLTDGLSSLR 348
QY 343 SETVKIHFSPVRPFTLP--GDE-----EESFIQIITPVR 376
Db 349 SERVSGFTTAGKALLRPVSGDDRPVAGLNGNGPFPAPVSTDYVYLLMPVR 399

RESULT 35

US-10-476-597-143
; Sequence 143, Application US/10476597
; Publication No. US20040235766A1
; GENERAL INFORMATION:
; APPLICANT: Bullard, James
; APPLICANT: Janjic, Nebojsa
; APPLICANT: McHenry, Charles S.
; TITLE OF INVENTION: System for Discovery of Agents that Block Yersinia Pestis and
; FILE REFERENCE: RDYN.02/PCT-US
; CURRENT APPLICATION NUMBER: US/10/476,597
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: US 60/290,725
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: PCT/US02/15111
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 60/332,644
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 143
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis

US-10-476-597-143
Query Match 19.5%; Score 364; DB 17; Length 402;
Best Local Similarity 25.8%; Pred. No. 1.7e-21;
Matches 106; Conservative 82; Mismatches 163; Indels 60; Gaps 11;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSIKIEVTSTGVTLTGSGNQISIENTIPVS 61
Db 13 LTRFLLRESFADAVSWAKNLPARPAPVLSGLVLTGSDNGLTISGFYEVSAEAQV--- 69
QY 62 NENAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGKQVD 121
Db 70 ----GAEIVSPGSLVSGRLSDITRALPNKQVDV-HVEGNRVALTTCGNARFSLTMPVE 124
QY 122 QYPRLOEVSTENPLIKTKLLKSIITAEATAFAASLOESRPLTGVHVLNHNKDKFVATD 181
Db 125 DYPTLTPTPEETGL-LPAELFAEASQVAIAAGRDITLPMLTGIRVILGETVWLA-ATD 182
QY 182 SHRMSQRLITLNTSADL--MVLPSKSLREFS-----AVFTDDIETV 222
Db 183 RFLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAGIGSDVRLSLGTCGPGVKDGLGI 242
QY 223 EVFFSPSQILFRSEHISFYTRLEGNYPDTDRLLMTETETEVVFNQSLRHAMERAFILIS 282

Db 243 SGNKRST-----TRLLDAEFKFRQLLPTEHTAVATMDVABIEAIAKLVALV- 290
QY 283 NATONGTVKLEITQNHISAHVNSPEVGKVNEDLDIVSOGSDLTISFNPTYLIESLKAIK 342
Db 291 -ADGAQVRMEFADGSRVLSAGADVDGRAEBDL-VVDYAGEPLTIAFNPTYLTDGLSSLR 348
QY 343 SETVKIHFSPVRPFTLP--GDE-----EESFIQIITPVR 376
Db 349 SERVSGFTTAGKALLRPVSGDDRPVAGLNGNGPFPAPVSTDYVYLLMPVR 399

RESULT 36

US-10-282-122A-62773
; Sequence 62773, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62773
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Mycobacterium bovis

US-10-282-122A-62773
Query Match 19.5%; Score 363; DB 15; Length 402;
Best Local Similarity 25.8%; Pred. No. 2e-21;
Matches 106; Conservative 81; Mismatches 164; Indels 60; Gaps 11;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSIKIEVTSTGVTLTGSGNQISIENTIPVS 61
Db 13 LTRFLLRESFADAVSWAKNLPARPAPVLSGLVLTGSDNGLTISGFYEVSAEAQV--- 69
QY 62 NENAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGKQVD 121
Db 70 ----GAEIVSPGSLVSGRLSDITRALPNKQVDV-HVEGNRVALTTCGNARFSLTMPVE 124

QY 122 QYPRLOEVSTENPLILKTKLLKSIITAETAFASQESRPILTGWHVLSNHHKDFKAVTD 191
 DB 125 DYPTLPTLPETGL-LPAELFAEASISQVATAAGDDTLPLMTGIRVELGETVVLAA-ATD 192
 QY 182 SHRMSQRLITLNTSADL--MVLPKSLREFS-----AVFTDDIETV 222
 DB 183 RFLAVRELKWSASSPDIEAALVPAKTLAAAKAGIGSDVRLSLGTGPGVKGDLGI 242
 QY 223 EVFSPQILFRSEHISFYTRLLEGNYPDTRLLMTFETEVVNTQSLRHAMERAFILIS 282
 DB 243 SNGKRT-----TRLDAEPKPRQLLPTEHTAVAMDAEILAEIKUVALV- 290
 QY 283 NATQNGTVKLEITQNHISAHVNSPEVGNVEDLIVSQSGDLTISFNPTYLIESLKAJK 342
 DB 291 -ADRGAVRMEFADGVSRLSAGADDVGRABEDL-VVDYAGEPLTIAFNPTYLTDGLSSLR 348
 QY 343 SETVKIHLSFVRPFTLTP--GDE-----EESFIQITPVR 376
 DB 349 SERVSFGFTTAGKPAALLRPVSGDDRPVAGLNGNGPFPFAVSTDYVYLLMPVR 399

RESULT 37
 US-10-282-122A-68144
 ; Sequence 68144, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 68144
 ; LENGTH: 367
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas putida
 US-10-282-122A-68144

Query Match 19.5%; Score 362.5; DB 15; Length 367;

Best Local Similarity 25.3%; Pred. No. 1.9e-21;
 Matches 97; Conservative 96; Mismatches 164; Indels 27; Gaps 11;
 QY 2 IQFSINRRLFHALNTTKRAISTKNAIPILSIIKIEVTSTGVTLTSGNSQISIENTIPVS 61
 DB 1 MHFTIQREALLKPLQLVAGVWERRQTLFVLSNVLLVVOGQQLSLTGTDELEVELVGRVQLE 60
 QY 62 NENAGLLTSPQAILLEASFFINIISLP-DISINVKIEIOHQVVLTSKGSEITLKGKDV 120
 DB 61 EP-----AEPGEITVPARKLMDICKSLPNDALIDIK-VDEQKLVKAGRSRFTLSLTPA 113
 QY 121 DOYPRLOEVSTENPLILKTKLLKSIITAETAFASQESRPILTGWHVLSNHHKDFKAVAT 180
 DB 114 NDFPVEGSPGLTCLNLEQSKRLRIERTSPAMAQQOVRYVINGMLLEVSRL-TLRVST 172
 QY 181 DSHRMSQRLITLNTSADL-----MVLPKSLREFSAVFTDDIETVEVFPSPQILFR 234
 DB 173 DGHRLA-----LCSMSAPIEQEDRHQVIVPRKGIELARLLTDPEGWMSIVLGQHHIRAT 227
 QY 235 SEHISFYTRLLEGNYPDTRLLMTFETEVVNTQSLRHAMERAFILISNATQNGTVKLEI 294
 DB 228 TGEFTTSLKLVGKFPDYPYERVLPGKGDKLVVGDQALREAFSRTAILSNEKYRG-IRLQL 286
 QY 295 TQNHISAHVNSPEVGNVEDLIVSQSGDLTISFNPTYLIESLKAJKSETVKIHLSFVSPV 354
 DB 287 AAGQIKIANNPEQEEAEIEIS-VDEGSSLEIGFVNSYLLDLVGLVMTTEQVRL-ILSDS 344
 QY 355 RPFTL--TPGDEESFQITPVR 376
 DB 345 NSSALLOEAGNDSSYV--VMPMR 366

RESULT 38
 US-10-671-403-112
 ; Sequence 112, Application US/10671403
 ; Publication No. US20040038289A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Yuzhakov, Alexander
 ; APPLICANT: Yurieva, Olga
 ; APPLICANT: Jeruzalmi, David
 ; APPLICANT: Bruck, Irina
 ; APPLICANT: Kuriyan, John
 ; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
 ; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
 ; FILE REFERENCE: 22221/1030
 ; CURRENT APPLICATION NUMBER: US/10/671,403
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US/09/716,964A
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416
 ; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 112
 ; LENGTH: 367
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas putida
 US-10-671-403-112

Query Match 19.5%; Score 362.5; DB 15; Length 367;
 Best Local Similarity 25.3%; Pred. No. 1.9e-21;
 Matches 97; Conservative 96; Mismatches 164; Indels 27; Gaps 11;
 QY 2 IQFSINRRLFHALNTTKRAISTKNAIPILSIIKIEVTSTGVTLTSGNSQISIENTIPVS 61
 DB 1 MHFTIQREALLKPLQLVAGVWERRQTLFVLSNVLLVVOGQQLSLTGTDELEVELVGRVQLE 60

QY 62 NENAGLLITSPGAILLEASFFINIISLP-DISINVKIEHQVLTSGKSEITLKGKDV 120
Db 61 EP-----AEPGEITVPARKMDICKSLPNDALIDIK-VDEQKLLVKAGRSRFTLTSLPA 113
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFASIQESRPILTGVHIVLSNKHDFKAVAT 180
Db 114 NDFPVEEGPGSLTNCLEQSKRLRIERTSFAMAQDVRYIYNGMLLEVSRLN-TLRAVST 172
QY 181 DSHRMSORLITLNTSADL-----MVLPSKSLREFSAVFTDDIETVEVFPSPQILFR 234
Db 173 DGHRLA-----LCSMSAPIEQDRHQVIVPRKGILELARLLTDPGVMVSVILGQHHIRAT 227
QY 235 SEHISFYTRLLEGNYPDTRLMTFETEVVNTQSLRHAMERAFILSNATQNGTVKLEI 294
Db 228 TGEFTFTSKLVGKFPDYERLPKGGDKLVGDRQALREAFRTAILSNEKVRG-IRLQ 286
QY 295 TQNHISAHVNSPEVGNEDLDIVSQSGSDLTISFNPTYLIESLKAISKSETVKIHLSPV 354
Db 287 AAGQLKIANNPEQEAEEIS-VDYEGSLEIGFNVSVLLDVLGVMTEQVRL-ILSDS 344
QY 355 RPFTL--TPGDEERSFIOLITPVR 376
Db 345 NSSALLOEAGNDSSYV--VMPMR 366

RESULT 39

US-10-671-419-112
; Sequence 112, Application US/10671419
; Publication No. US20040038290A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 2221/1030
; CURRENT APPLICATION NUMBER: US/10/671,419
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 112
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-671-419-112

Query Match 19.5%; Score 362.5; DB 15; Length 367;
Best Local Similarity 25.3%; Pred. No. 1.9e-21;
Matches 97; Conservative 96; Mismatches 164; Indels 27; Gaps 11;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSIKIEVTSTGVTGSGNQISIENTIPVS 61
Db 1 MHFTIQRALLKPLQVAGVVERQTLPVLSNVLLVQSQQLSLGTDLLEVELVGRVQLE 60
QY 62 NENAGLLITSPGAILLEASFFINIISLP-DISINVKIEHQVLTSGKSEITLKGKDV 120
Db 61 EP-----AEPGEITVPARKMDICKSLPNDALIDIK-VDEQKLLVKAGRSRFTLTSLPA 113
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFASIQESRPILTGVHIVLSNKHDFKAVAT 180
Db 114 NDFPVEEGPGSLTNCLEQSKRLRIERTSFAMAQDVRYIYNGMLLEVSRLN-TLRAVST 172

QY 181 DSHRMSORLITLNTSADL-----MVLPSKSLREFSAVFTDDIETVEVFPSPQILFR 234
Db 173 DGHRLA-----LCSMSAPIEQDRHQVIVPRKGILELARLLTDPGVMVSVILGQHHIRAT 227
QY 235 SEHISFYTRLLEGNYPDTRLMTFETEVVNTQSLRHAMERAFILSNATQNGTVKLEI 294
Db 228 TGEFTFTSKLVGKFPDYERLPKGGDKLVGDRQALREAFRTAILSNEKVRG-IRLQ 286
QY 295 TQNHISAHVNSPEVGNEDLDIVSQSGSDLTISFNPTYLIESLKAISKSETVKIHLSPV 354
Db 287 AAGQLKIANNPEQEAEEIS-VDYEGSLEIGFNVSVLLDVLGVMTEQVRL-ILSDS 344
QY 355 RPFTL--TPGDEERSFIOLITPVR 376
Db 345 NSSALLOEAGNDSSYV--VMPMR 366

RESULT 40

US-10-670-844-112
; Sequence 112, Application US/10670844
; Publication No. US20040043414A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Yuzhakov, Alexander
; APPLICANT: Yurieva, Olga
; APPLICANT: Jeruzalmi, David
; APPLICANT: Bruck, Irina
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 2221/1030
; CURRENT APPLICATION NUMBER: US/10/670,844
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US/09/716,964A
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/143,202
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 08/823,407
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 09/057,416
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 112
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-670-844-112

Query Match 19.5%; Score 362.5; DB 15; Length 367;
Best Local Similarity 25.3%; Pred. No. 1.9e-21;
Matches 97; Conservative 96; Mismatches 164; Indels 27; Gaps 11;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSIKIEVTSTGVTGSGNQISIENTIPVS 61
Db 1 MHFTIQRALLKPLQVAGVVERQTLPVLSNVLLVQSQQLSLGTDLLEVELVGRVQLE 60
QY 62 NENAGLLITSPGAILLEASFFINIISLP-DISINVKIEHQVLTSGKSEITLKGKDV 120
Db 61 EP-----AEPGEITVPARKMDICKSLPNDALIDIK-VDEQKLLVKAGRSRFTLTSLPA 113
QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFASIQESRPILTGVHIVLSNKHDFKAVAT 180
Db 114 NDFPVEEGPGSLTNCLEQSKRLRIERTSFAMAQDVRYIYNGMLLEVSRLN-TLRAVST 172
QY 181 DSHRMSORLITLNTSADL-----MVLPSKSLREFSAVFTDDIETVEVFPSPQILFR 234
Db 173 DGHRLA-----LCSMSAPIEQDRHQVIVPRKGILELARLLTDPGVMVSVILGQHHIRAT 227
QY 235 SEHISFYTRLLEGNYPDTRLMTFETEVVNTQSLRHAMERAFILSNATQNGTVKLEI 294

Db	228	TGEFTTSKLVGKFPDYERVLPGGDKLVVGDQALREAFRTAILSNKRYG-IRLQL	286
Qy	295	TONHISAHVNSPEYGVKYNEDLDIVSQSGDLTISFNFTYLIESLKAIKSETVKIHFSPV	354
Db	287	AAGQIKIQANNPEOEAEIEIS-VDYEGSSLEIGFNVSYLLDLVGVMTTEQVRL-IISDS	344
Qy	355	RPFTL--TPGDEESFIQLITPVR	376
Db	345	NSSALLOEAGNDSSYV--VMFMR	366

Search completed: January 28, 2005, 17:12:11
 Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2005, 16:35:51 ; Search time 157 Seconds
(without alignments)
863.692 Million cell updates/sec

Title: US-10-048-071-28
Perfect score: 1863
Sequence: 1 MIQFSINRFLFIHALNTTKR.....LTPGDSEBSFIQLITPVRTN 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1863	100.0	378	4 AAB31944	Aab31944 Amino aci
2	1849	99.2	378	5 ABP27970	Abp27970 Streptoco
3	1849	99.2	378	5 ABG61623	Abg61623 Streptoco
4	1849	99.2	378	6 ABU46377	Abu46377 Protein e
5	1527	82.0	378	6 ABU44094	Abu44094 Protein e
6	1491	80.0	378	5 ABP27969	Abp27969 Streptoco
7	1389	74.6	378	6 ABU02772	Abu02772 S. pneumo
8	1389	74.6	378	8 ADK47890	Adk47890 Streptoco
9	1384	74.3	378	4 AAU37853	Aau37853 Streptoco
10	1384	74.3	378	6 ABU45758	Abu45758 Protein e
11	1053	56.5	380	5 ABB53301	Abb53301 Lactococc
12	864.5	46.4	376	4 AAU35314	Aau35314 Enterococ
13	864.5	46.4	376	6 ABU29060	Abu29060 Protein e
14	863.5	46.3	384	7 ADC94929	Adc94929 E. faeciu
15	862.5	46.3	376	6 ABU30083	Abu30083 Protein e
16	748.5	40.2	377	6 ABU42909	Abu42909 Protein e
17	748.5	40.2	385	5 ABP40321	Abp40321 Staphyloc
18	745	40.0	400	6 ADB07184	Adb07184 Alloiocec
19	745	40.0	400	8 ADJ27077	Adj27077 Alloiocec
20	735.5	39.5	377	6 ABU43984	Abu43984 Protein e
21	730.5	39.2	377	2 AAU49071	Aau49071 Amino aci
22	730.5	39.2	377	4 AAU37133	Aau37133 Staphyloc
23	730.5	39.2	377	4 AAU37107	Aau37107 Staphyloc
24	730.5	39.2	377	4 AAB31935	Aab31935 Amino aci
25	730.5	39.2	377	5 ABB81847	Abb81847 Staphyloc

ALIGNMENTS

RESULT 1
AAB31944

ID AAB31944 standard; protein; 378 AA.

AC AAB31944;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a partial dnaN (beta subunit) polypeptide.

KW dnaG; Gram positive bacteria; polC; dnaE; hola; holB; dnaX; dnaN; seb;
KW dnaG; dnaB; antibiotic; replication; cell growth; cell death;
KW bacterial infection.

OS Streptococcus pyogenes.

PN WO200109164-A2.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US020666.

PR 29-JUL-1999; 99US-0146178P.

XX (UVRQ) UNIV ROCKEFELLER.

O'donnell ME, Bruck I, Zhang D, Whipple R;

WPI; 2001-147453/15.

N-PSDB; AAF54744.

Isolated DNA molecule from a Gram positive bacterium encoding DNA replication proteins used to identify compounds which have antibiotic activity.

Claim 80; Page 62-63; 239pp; English.

The present sequence represents a partial dnaN polypeptide. The specification describes DNA molecules from Gram positive bacteria, which comprise a coding region from a polC, dnaE, hola, holB, dnaX, dnaN, seb, dnaG or a dnaB gene. These sequences encode proteins that replicate the chromosome of Gram positive bacteria. They are used for sequencing and amplification of DNA and in drug discovery to identify compounds which have antibiotic activity through interference with replication. They are used in methods for identifying compounds that are active at the level of DNA replication and result in arrest of cell growth or cell death of bacteria to treat bacterial infections in animals

XX

SQ Sequence 378 AA;

Query Match 100.0%; Score 1863; DB 4; Length 378;
 Best Local Similarity 100.0%; Pred. No. 7.1e-156;
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
 DB 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
 QY 61 SNEAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGDV 120
 DB 61 SNEAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGDV 120
 QY 121 DQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTCVHIVLSNKHDFKAVAT 180
 DB 121 DQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTCVHIVLSNKHDFKAVAT 180
 QY 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 DB 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 QY 241 YTRLLEGNYPTDRLMTTEFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNHIS 300
 DB 241 YTRLLEGNYPTDRLMTTEFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNHIS 300
 QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVPRPFTLT 360
 DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVPRPFTLT 360
 QY 361 PGDEESFIQLITPVRTN 378
 DB 361 PGDEESFIQLITPVRTN 378

RESULT 2
 ABP27970
 ID ABP27970 standard; protein; 378 AA.

XX AC ABP27970;

DT 02-JUL-2002 (first entry)

XX STREPTOCOCCUS POLYPEPTIDE SEQ ID NO 5116.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

XX 24-NOV-2000; 2000GB-00028727.

XX 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;

XX Tettelin H;

XX WPI; 2002-352536/38.

XX N-PSDB; AEN68601.

XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.

XX PS

Claim 1; Page 3675; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus (GBS) (Streptococcus agalactiae) or group A streptococcus (GAS) (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins

SQ Sequence 378 AA;

Query Match 99.2%; Score 1849; DB 5; Length 378;
 Best Local Similarity 99.2%; Pred. No. 1.2e-154;
 Matches 375; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
 DB 1 MIQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGQISIENTIPV 60
 QY 61 SNEAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGDV 120
 DB 61 SNEAGLLITSPGAILLEASFFNIISLPDISINVKIEHQHVLTSGKSEITLKGDV 120
 QY 121 DQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTCVHIVLSNKHDFKAVAT 180
 DB 121 DQYPRLOEVSTENPLILKTKLSIIAETAFASLOESRPILTCVHIVLSNKHDFKAVAT 180
 QY 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 DB 181 DSHRMSQRLITLNDTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
 QY 241 YTRLLEGNYPTDRLMTTEFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNHIS 300
 DB 241 YTRLLEGNYPTDRLMTTEFETEVNTQSLRHAMERAFILSNATONGTVKLEITQNHIS 300
 QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVPRPFTLT 360
 DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAIKSETVKIHFLSPVPRPFTLT 360
 QY 361 PGDEESFIQLITPVRTN 378
 DB 361 PGDEESFIQLITPVRTN 378

RESULT 3

ABG61623

ID ABG61623 standard; protein; 378 AA.

XX AC ABG61623;

XX 12-AUG-2002 (first entry)

XX Streptococcus pyogenes DnaN protein beta sub-unit.

XX DNA polymerase III holoenzyme subunit protein; replicase; polymerase;
 KW antibacterial; vaccine.

XX Streptococcus pyogenes.

XX WO200234936-A2.


```
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US048396.
XX
PR 27-OCT-2000; 2000US-0244023P.
XX
PA (REPL-) REPLIDYNE INC.
XX
PI Mchenry CS, Bullard JM, Janjic N, Manhardt EL, Kery V;
PI Williams JC;
XX
DR WPI; 2002-463366/49.
XX
DR N-PSDB; ABK83362, ABK83362.
XX
PT Novel bacterial DNA polymerase III subunit proteins and genes encoding
PT the proteins, for reconstituting replicases and polymerases for
PT sequencing, amplification and screening compounds that modulate
PT polymerase.
XX
PS Claim 9; Fig 21B; 268pp; English.
XX
CC The invention describes an isolated bacterial, preferably Streptococcus
CC pyogenes DNA polymerase III holoenzyme subunit proteins (I). The protein
CC and the polynucleotide encoding it are useful for reconstituting
CC replicases and polymerases for sequencing, amplification and screening
CC for compounds which modulate the function of polymerase or replicase. An
CC antibody to the protein is useful as therapeutic compound to passively
CC immunise an animal to protect the animal from bacteria susceptible to
CC treatment by such antibodies, preferably Streptococcus pyogenes, as
CC reagents in assays to detect infection by such bacteria, and/or as tools
CC to screen expression libraries and/or to recover desired proteins from a
CC mixture of protein and other contaminants. This is the amino acid
CC sequence of a Streptococcus pyogenes polymerase III holoenzyme sub-unit
XX
SQ Sequence 378 AA;

Query Match          99.2%; Score 1849; DB 5; Length 378;
Best Local Similarity 99.2%; Pred. No. 1.2e-154;
Matches 375; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIQFSINTLFLHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGQISIENTIPV 60
DB 1 MIQFSINTLFLHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGQISIENTIPV 60

QY 61 SNEAGLLITSGAILLEASPFNIISLDPDISINVKEIQHVVLTSKSEITLKGKDV 120
DB 61 SNEAGLLITSGAILLEASPFNIISLDPDISINVKEIQHVVLTSKSEITLKGKDV 120

QY 121 DOYPRLOEVSTENPLILTKLKSIIAETAFASLOESRPILTGVHIVLSNKHDKFVAT 180
DB 121 DOYPRLOEVSTENPLILTKLKSIIAETAFASLOESRPILTGVHIVLSNKHDKFVAT 180

QY 181 DSHRMSQRLITLDNTSADLMVLPKSLREFSAVFTDDIETVEVFPQSILFRSEHSIF 240
DB 181 DSHRMSQRLITLDNTSADLVVLPKSLREFSAVFTDDIETVEVFPQSILFRSEHSIF 240

QY 241 YTRLLEGNYPDTDRLLMTFETFEVFNFTQSLRHAMERAFLLSNATQNGTVKLEITQNHIS 300
DB 241 YTRLLEGNYPDTDRLLMTFETFEVFNFTQSLRHAMERAFLLSNATQNGTVKLEITQNHIS 300

QY 301 AHVNSPEVGNKVEDIDIVSQSGSDLTISFNPTYLIESLKAISKSEIVKIHFLSPVPFPLT 360
DB 301 AHVNSPEVGNKVEDIDIVSQSGSDLTISFNPTYLIESLKAISKSEIVKIHFLSPVPFPLT 360

QY 361 PGDEESFQLITPVRTN 378
DB 361 PGDEESFQLITPVRTN 378

RESULT 4
ABU46377
ID ABU46377 standard; protein; 378 AA.
XX
```

```
AC ABU46377;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #31904.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Streptococcus pyogenes.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR N-PSDB; ACA50247.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 74301; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 378 AA;

Query Match          99.2%; Score 1849; DB 6; Length 378;
Best Local Similarity 99.2%; Pred. No. 1.2e-154;
Matches 375; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 MIQFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
DB 1 MIQFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
QY 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
DB 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120
QY 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNKHDKFVAVT 180
DB 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNKHDKFVAVT 180
QY 181 DSHRMSORLITLNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
DB 181 DSHRMSORLITLNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240
QY 241 YTRLLEGNYPDRLMLTTEFEVVTQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
DB 241 YTRLLEGNYPDRLMLTTEFEVVTQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTVLTIESLKAISKSEITVKHFLSPVRPFTLT 360
DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTVLTIESLKAISKSEITVKHFLSPVRPFTLT 360
QY 361 PGDEESFIQLITPVRTN 378
DB 361 PGDEESFIQLITPVRTN 378

RESULT 5

ABU44094

ID ABU44094 standard; protein; 378 AA.

AC ABU44094;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #29621.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Streptococcus mutans.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-029926/02.

XX N-PSDB; ACA47964.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 72018; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 378 AA;

Query Match 82.0%; Score 1527; DB 6; Length 378;

Best Local Similarity 79.9%; Pred. No. 3.7e-126;

Matches 302; Conservative 39; Mismatches 37; Indels 0; Gaps 0;

QY 1 MIQFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60

DB 1 MIQFSINRTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60

QY 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120

DB 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEHQVVLTSKSEITLKGKDV 120

QY 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNKHDKFVAVT 180

DB 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNKHDKFVAVT 180

QY 181 DSHRMSORLITLNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240

DB 181 DSHRMSORLITLNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHISF 240

QY 241 YTRLLEGNYPDRLMLTTEFEVVTQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300

DB 241 YTRLLEGNYPDRLMLTTEFEVVTQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300

QY 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTVLTIESLKAISKSEITVKHFLSPVRPFTLT 360

DB 301 AHVNSPEVGKVNEDLDIVSQSGDLTISFNPTVLTIESLKAISKSEITVKHFLSPVRPFTLT 360

QY 361 PGDEESFIQLITPVRTN 378

DB 361 PGDEESFIQLITPVRTN 378

RESULT 6

ABP27969

ID ABP27969 standard; protein; 378 AA.

XX ABP27969;

XX AC

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 5114.

CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 CC
 XX

SQ Sequence 378 AA;

Query Match 74.6%; Score 1389; DB 6; Length 378;
 Best Local Similarity 72.2%; Pred. No. 6e-114;
 Matches 273; Conservative 50; Mismatches 55; Indels 0; Gaps 0;

QY 1 MIOFSINRTLFHAIANTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
 Db 1 MIHFSINKNLFQALNTTKRAISSKNAIPILSTVKIDVTNEGITLIGSNGQISIENTISQ 60
 QY 61 SNENAGLLITSPGAILLEASFFINIISSLPDISINVKIEHQVVLTSKSEITLKGKDV 120
 Db 61 KNEDAGLLITSLGAILLEASFFINNVSSLPDVLDFKEIQNVLTSGKSEITLKGKDS 120
 QY 121 DOYPRLOEVSTENPILKTLKLSIAETAFAASIQESRPILTGVIHVLNHNKOPKAVAT 180
 Db 121 EQYPRIOEISASTPLILETKLKKIINETAFAASTQESRPILTGVIHVLNHNKOPKAVAT 180
 QY 181 DSHRMSORLITLDMNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFREHSIF 240
 Db 181 DSHRLSOKKLTLEKNSDDFDVIPSRLREFSAVFTDDIETVEIFFANQLFRSENISF 240
 QY 241 YTRLLEGNYPDTDRLLMTFETEVVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
 Db 241 YTRLLEGNYPDTDRLLMTFETEVVFNQSLRHAMERAFILSNATQNGTVKLEIKDGVS 300
 QY 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLIESLKATSEVTKIHFILSPVRPFTLT 360
 Db 301 AHVHSPGVKNEEDTDQVTGEBDLTISFNPTLIDSLKALNSEKVTISFISAVRPFTLV 360
 QY 361 PGDEESFIQLITPVRTN 378
 Db 361 PADTDEDFMQLITPVRTN 378

RESULT 8

ADK47890
 ID ADK47890 standard; protein; 378 AA.

XX

AC ADK47890;

XX 20-MAY-2004 (first entry)

XX Streptococcus pneumoniae protein, Seq ID No 4405.

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

XX Streptococcus pneumoniae.

XX US6699703-B1.

XX 02-MAR-2004.

XX 26-MAY-2000; 2000US-00563110.

XX 02-JUL-1997; 97US-0051553P.

XX 12-MAY-1998; 98US-0085131P.

XX 30-JUN-1998; 98US-00107433.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;

XX

DR WPI; 2004-212399/20.
 DR N-PSDB; ADK45229.

XX New nucleic acid molecules and polypeptides useful for diagnosing,
 PT preventing and treating pathological conditions resulting from bacterial
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug
 PT screening.

XX Disclosure; SEQ ID NO 4405; 301pp; English.

XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
 CC and polypeptides. The nucleic acids and proteins are useful for
 CC diagnosing, preventing and treating pathological conditions resulting
 CC from bacterial infection, such as S. pneumoniae infection. These may also
 CC be used for drug screening procedures. The present sequence represents a
 CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
 CC data for this patent did not appear in the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html.

SQ Sequence 378 AA;

Query Match 74.6%; Score 1389; DB 8; Length 378;
 Best Local Similarity 72.2%; Pred. No. 6e-114;
 Matches 273; Conservative 50; Mismatches 55; Indels 0; Gaps 0;

QY 1 MIOFSINRTLFHAIANTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
 Db 1 MIHFSINKNLFQALNTTKRAISSKNAIPILSTVKIDVTNEGITLIGSNGQISIENTISQ 60
 QY 61 SNENAGLLITSPGAILLEASFFINIISSLPDISINVKIEHQVVLTSKSEITLKGKDV 120
 Db 61 KNEDAGLLITSLGAILLEASFFINNVSSLPDVLDFKEIQNVLTSGKSEITLKGKDS 120
 QY 121 DOYPRLOEVSTENPILKTLKLSIAETAFAASIQESRPILTGVIHVLNHNKOPKAVAT 180
 Db 121 EQYPRIOEISASTPLILETKLKKIINETAFAASTQESRPILTGVIHVLNHNKOPKAVAT 180
 QY 181 DSHRMSORLITLDMNTSADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFREHSIF 240
 Db 181 DSHRLSOKKLTLEKNSDDFDVIPSRLREFSAVFTDDIETVEIFFANQLFRSENISF 240
 QY 241 YTRLLEGNYPDTDRLLMTFETEVVFNQSLRHAMERAFILSNATQNGTVKLEITQNHIS 300
 Db 241 YTRLLEGNYPDTDRLLMTFETEVVFNQSLRHAMERAFILSNATQNGTVKLEIKDGVS 300
 QY 301 AHVNSPEVGVKNEDLDIVSQSGDLTISFNPTLIESLKATSEVTKIHFILSPVRPFTLT 360
 Db 301 AHVHSPGVKNEEDTDQVTGEBDLTISFNPTLIDSLKALNSEKVTISFISAVRPFTLV 360
 QY 361 PGDEESFIQLITPVRTN 378
 Db 361 PADTDEDFMQLITPVRTN 378

RESULT 9

AAU37853

ID AAU37853 standard; protein; 378 AA.

XX

AC AAU37853;

XX 14-FEB-2002 (first entry)

XX Streptococcus pneumoniae cellular proliferation protein #282.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;

XX antibacterial; drug design.

XX Streptococcus pneumoniae.

XX WO200170955-A2.

XX 27-SEP-2001.

```
XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207272P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlseen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX N-PSDB; NAA55712.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 13446; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes,
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 378 AA;
XX
XX Query Match 74.3%; Score 1384; DB 4; Length 378;
XX Best Local Similarity 72.2%; Pred. No. 1.6e-113;
XX Matches 273; Conservative 49; Mismatches 56; Indels 0; Gaps 0;
XX
XX 1 MIQPSINTPLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVTTGNGQISIENTIPV 60
XX 1 MIHFSINKNLFQALNITKRAISSKNAIPILSTKIDVTNEGTVLIGSNGQISIENTISQ 60
XX
XX 61 SNEAGLLITSPGAILLEAFSPINIISSLPDISINVKETEOHQVVLTSKSRITLKGKDV 120
XX 61 KNEAGLLITSGSILLESFNFVNSLDPDVLDFKEQIVLTSGKSRITLKGKDS 120
XX
XX 121 DQYPRLOEVSTENPLILTKLKLSIAETAFAASLQESRPILTGVHIVLSNHNKOPKAVAT 180
XX 121 EQYPRIQEISASTPLILETKLKLIINETAFAASTQESRPILTGVHIVLSQHKELKTAVT 180
XX
XX 181 DSHRMSQRILTDNTSADLMVLPKSLUREPSAVFTDDIETVEVFPSPQILFRSEHISF 240
XX 181 DSHRLSQKLTLEKNSDDPDVVPISRLSREFSAVFTDDIETVEIIFANNQILFRSENISF 240
XX
XX 241 YTRLLEGNYPDTRLIMTEFEVVFNTOSTLRHMERAPLISNATQNGTVKLEITQNHIS 300
XX 241 YTRLLEGNYPDTRLIPTDNTFTTFFNVNLRQSMERAPLSSATQNGTVKLEIKDGVVS 300
XX
XX 301 AHVNSPEVGNVEDLDIVSQSGDLTISFNPYLIIESKAIKSETVKIHFSPVPFFTLT 360
XX 301 AHVHSPVGVKNEEDTDQVGTEDLTISFNPYLIIDSLKALNSEKVTISFISAVPFFTLV 360
```

```
QY 361 PGDEERSFIQLITPVRTN 378
DB 361 PADTDEDFMQLITPVRTN 378
RESULT 10
ID ABU45758
XX ABU45758 standard; protein; 378 AA.
XX AC ABU45758;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #31285.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Streptococcus pneumoniae.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA49628.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 73682; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
```

CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 378 AA;

Query Match 74.3%; Score 1384; DB 6; Length 378;
 Best Local Similarity 72.2%; Pred. No. 1.6e-113;
 Matches 273; Conservative 49; Mismatches 56; Indels 0; Gaps 0;

QY 1 MIOFSINRFLFIHALNNTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
 DB 1 MIHFSINKNLFQALNITKRAISSKNAIPILSKIDVTNEGVTIGSNGQISIENTISQ 60
 QY 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEQHVLTSGKSEITLKGKDV 120
 DB 61 KNEADAGLLITSLGILLASFFINNVSSLPDVTLDKFEIQVLTSGKSEITLKGKDS 120
 QY 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNHNKDFKAVAT 180
 DB 121 EQYPRIOEISASPLILETKLKKIINETAFASLOESRPILTGVIHVLNHNKDFKAVAT 180
 QY 181 DSHRMSORLITLDMTADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHSF 240
 DB 181 DSHRLSQKLTLEKNSDDFDVIPSRLREFSAVFTDDIETVEIIPFANNQILFRSENISF 240
 QY 241 YTRLLEGNYPTDRLMTTEFETEVVFNTOQLRHAMERAPLISNATQNGTVKLEITQNHIS 300
 DB 241 YTRLLEGNYPTDRLIPTDFNTITFNVNVLQSMERARLLSSATQNGTVKLEIKDGVVS 300
 QY 301 AHVNSPEVCKNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSPVRPFTLT 360
 DB 301 AHVHSPGVKNEEDTDQVTGEDLTISFNPTYLIDSLKALNSEKVTISFISAVRPFTLV 360
 QY 361 PGDEESFQILITPVRTN 378
 DB 361 PADTDEDFQILITPVRTN 378

RESULT 11
 ABB53301
 ID ABB53301 standard; protein; 380 AA.
 XX
 AC ABB53301;
 XX
 DT 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein dnan.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis; IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-00004630.
 XX
 PR 11-APR-2000; 2000FR-00004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR WPI; 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species.
 XX
 PS Claim 6; SEQ ID NO 3; 2504pp; French.
 XX

CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO2001/77334 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 380 AA;

Query Match 56.5%; Score 1053; DB 5; Length 380;
 Best Local Similarity 54.7%; Pred. No. 3.1e-84;
 Matches 208; Conservative 73; Mismatches 97; Indels 2; Gaps 1;

QY 1 MIOFSINRFLFIHALNNTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
 DB 1 MIKFSINKNAPQNALRITKQAGSKVTIPALTKLIEVENGITLIGSNGQISIKNPLPV 60
 QY 61 SNEAGLLITSPGAILLEASFFINIISLPDISINVKIEQHVLTSGKSEITLKGKDV 120
 DB 61 DNKDASMLISCTGSLLEAFAFFENVVSQLPEVLEFTEKSKQVLLTSGKSEITLKGDS 120
 QY 121 DQYPRLOEVSTENPLILKTLKLSIIAETAFASLOESRPILTGVIHVLNHNKDFKAVAT 180
 DB 121 EYFPHLOEISEGSLKMKVKVLEIFETVFAVSTQENRPITFTGVHLETLTSGELKAVAT 180
 QY 181 DSHRMSORLITLDMTADLMVLPKSLREFSAVFTDDIETVEVFPSPQILFRSEHSF 240
 DB 181 DSHRMSORLPLESEKLFVILPSKINSFKVFTNDEEIEIFISGQMLFRNETISY 240
 QY 241 YTRLLEGNYPTDRLMTTE--FETEVVFNTOQLRHAMERAPLISNATQNGTVKLEITQNH 298
 DB 241 YSRLIEGSPDTRNLIPEADYTLDLVFDAAQLRHTMDRALLTVMTTNGTVKLTSGDS 300
 QY 299 ISAHVNSPEVCKNEDLDIVSQSGDLTISFNPTYLIESLKAISKSETVKIHFSPVRPPT 358
 DB 301 VVTTANSPEVSGVHEELTALSKEGNDLAIISFNPEYLDALKVKAPEVIRFISNVRPPT 360
 QY 359 LTPGDEESFQILITPVRTN 378
 DB 361 LQPNESGFGVQILITPVRTN 380

RESULT 12
 AAU35314
 ID AAU35314 standard; protein; 376 AA.
 XX
 AC AAU35314;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Enterococcus faecalis cellular proliferation protein #601.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS53173.
 XX
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 XX
 PS Example 3; SEQ ID NO 10907; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 376 AA;

ID ABU29060 standard; protein; 376 AA.
 AC ABU29060;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #14587.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA32930.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 56984; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 376 AA;

RESULT 13

ABU29060

Query Match

46.4%; Score 864.5; DB 6; Length 376;

[illegible]

CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)
xx
xx Sequence 377 AA;
SQ

Query Match	40.2%;	Score	748.5;	DB	6;	Length	377;	
Best Local Similarity	40.5%;	Pred. No.	2.7e-57;					
Matches	153;	Conservative	86;	Mismatches	136;	Indels	3;	
							Gaps	3;
QY	1	MIQFSNRTLFIHAHNTTKAISTKNAIPILSSIKIEVTSVGTVLTGSGNQISIENTIPV	60					
Db	1	MMSEFTIKRDFINGQNTLKAISPTTLPITGKIDAKENEVLTGSDSEISIEITPK	60					
QY	61	SNENAGLL-ITSPGAILLEASFFINIILSSIPDISINVKEIEHQHVLTGSKSEITLTKGD	119					
Db	61	QVQDEEIVEITGVSVPGLGRFFVDIIKLPGRKVLSTNEQFQTLITSGHSEFNLGSD	120					
QY	120	VQQYPRLOEYSTENPLIKTKLKSIIAETAFAASLOESRPLITGVHIVSNHKDFKAVA	179					
Db	121	PQQYPLLPVEVRSDAIQSVKVLKNIITAOQNFVSTSETPVLTVGNVLIQDN-ELICTA	179					
QY	180	TDSHRMSQRILITDNTSADIMVLPVLSKSLREFSAVFTDDIETVEVPFSPQILPRSHIS	239					
Db	180	TDSHRLAVRKLQLEDESENKNVILPGKALSELKIMSDSEDDIDIFFASQVILFRVGNIN	239					
QY	240	FYTRLLEGNPDPDRLLMTFETFEVNTOSLRHAMERAFILSNATQGTVKLEITQNH	299					
Db	240	FISRLLEGHPDTRLPFENYEIKLGINNGDFYHAIDRASILLAREGNNVILKSTGNELV	299					
QY	300	SAHWNGPEVKVNEIDLIVSQSGDLTISFNPPYLIESLKAISKSETVKIHLFSVPFPFL	359					
Db	300	ELSTSPFEITVKEEVNANDVEGNLKLISFNKYMMDKALKAINDVEVEFFGTMTKPFIL	359					
QY	360	TPGDEESFTQILTPVRT	377					
Db	360	KPKD-DDSVTQLILPRT	376					

RESULT 17
ABP40321
ID ABP40321 standard; protein: 385 AA.

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5166.

Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
antibacterial; gene therapy.

Staphylococcus epidermidis

PN US6380370-B1.

PD 30-APR-2002.

13-AUG-1998; 98US-00134001.

PR 14-AUG-1997; 97US-0055779P.

PR 08-NOV-1997; 97US-0064964P.
XX

PA (GENO-) GENOME THERAPEUTICS CORP.
YY

PI Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.
XX N-PSDB; ABN92866.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
XX polypeptide, useful for diagnosing and treating bacterial infections.
XX
XX Disclosure; SEQ ID NO 5166; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences can
XX also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life cycle
XX or inhibit S. epidermidis infection. N.B. The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
XX
XX Sequence 385 AA;
XX

[illegible]

RESULT 18
ADB07184

ID ADB07184 standard; protein; 400 AA.

AC ADB07184;

DT 20-NOV-2003 (first entry)

DE Alloiococcus otitis antigenic protein SEQ ID NO:1124.

Alloicoccus otitidis: antigenic protein; immunogenic; immunisation; gene therapy; Gram-positive bacterium; infection.

OS Alloiococcus otitis.

XX
PN WO2003048304-A2.XX
PD 12-JUN-2003.

```

XX PF 25-NOV-2002; 2002WO-US036123.
XX PR 29-NOV-2001; 2001US-0333777P.
XX PR 18-NOV-2002; 2002US-0426742P.
XX PA (AMHP ) WYETH HOLDINGS CORP.
XX PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;
XX DR WPI; 2003-505284/47.
XX DR N-PSDB; ADB07183.
XX PT New Alloiooccus otitidis polynucleotides and polypeptides, useful for
XX PT treating and diagnosing diseases, drug screening assays and monitoring of
XX PT effects during drug clinical trials.
XX PS Claim 33; SEQ ID NO 1124; 1019pp; English.
XX CC The present invention describes an isolated polynucleotide (I) of
XX CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
XX CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
XX CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
XX CC expression vector comprising the novel isolated polynucleotide (I); its
XX CC complement, degenerate variant or fragment; (3) a genetically engineered
XX CC host cell, transfected, transformed or infected with the vector of (2);
XX CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX CC composition comprising the polypeptide, its complement, biological
XX CC equivalent or fragment, or the polynucleotide that is comprised in the
XX CC expression vector; (6) a pharmaceutical composition comprising the
XX CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX CC of the polypeptides of (1), their biological equivalent or fragment; (8)
XX CC immunising against Alloiooccus otitidis by administering to a host the
XX CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
XX CC otitidis in the biological sample; (10) a kit comprising a container
XX CC containing the novel polynucleotide, its degenerate variant or fragment,
XX CC or the antibody of (4); and (11) producing a polypeptide by culturing the
XX CC genetically engineered host cell under conditions suitable to produce the
XX CC polypeptide from the culture. (I) can be used in gene therapy. The
XX CC polynucleotides, polypeptides, antibodies and compositions of the present
XX CC invention can be used for treating and diagnosing diseases, drug
XX CC screening assays and monitoring of effects during drug clinical trials.
XX CC The polynucleotides are useful for expressing and detecting Alloiooccus
XX CC otitidis. The present sequence represents an Alloiooccus otitidis
XX CC antigen protein from the present invention.
XX SQ Sequence 400 AA;

Query Match 40.0%; Score 745; DB 6; Length 400;
Best Local Similarity 42.3%; Pred. No. 6e-57;
Matches 160; Conservative 68; Mismatches 146; Indels 4; Gaps 4;

QY 2 1QFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVSTGVTLTGSNGQISIENTIPVS 61
DB 24 LKFSVKRTFELKVLKKVQIAVSSKSTIAITLKIEADNQGLTIGSNDSDISVESYLSVT 83
QY 62 NENAGLLITSPGAILLE-ASFFINIISSLPDISINVKIEHQHVLTSGKSEITLKGXDV 120
DB 84 DEGADLVIDEPGQIVLPQARLFANIVOKLPDTHFKVNVSSQGOQTQITSASASFTINGIDA 143
QY 121 DOYPLQEVSTENPLILTKLLKLSIAETAFAASLOESRPILTGVHIVLSNKHDPKAVAT 180
DB 144 MSYPLHPDIDLESFTPLPVDLFPKMINQTVIAVSNHESRPILTVGNLSLKEGR-LKAVAT 202
QY 181 DSHRMSQRLITLDNT-SADLMVVLPSKSLREPSAVFTDDIETVEFFPSQILFRSEHIS 239
DB 203 DSHRLSQRSIQLESAPDISFDIVIPGKSLTETLKMAD-EEVRVAISDNQILFALSQQ 261
QY 240 FYTRLEGNYPDTRLLMTEFEVFNQTSRSHAMERAFILSNATQNGTVKLEITQNH 299
DB 262 FYSRLLEGKYPTDRLIPGDTPTETITLDTKELQGVADRASLLSHEGKNVVLQTVTAEL 321
QY 300 SAHVNSPEVGKVNEDLDIVSSQGSDLTISFNPTYLIESLKAIKSETVKIHLFSPVRPTL 359

```

322 BIEGSAEVGHVQBEIDFGHFQGDLLISFPDYLKEALATFGQGEIKLVLSTLRPFVI 381

360 TPGDEERSFIQLITPVRT 377

382 VPSEDQGDIFIQLITPIRT 399

RESULT 19

ADJ27077

ID ADJ27077 standard; protein; 400 AA.

XX AC ADJ27077;

XX DT 20-MAY-2004 (first entry)

XX DE Alloiooccus otitidis protein, a novel antibacterial target SeqID 22.

XX KW growth and survival; otitis media with effusion; OME;

XX KW bacterial infection; antibacterial; antineurotic.

XX OS Alloiooccus otitis.

XX PN WO2003104391-A2.

XX PD 18-DEC-2003.

XX PF 25-NOV-2002; 2002WO-US036122.

XX PR 29-NOV-2001; 2001US-0333777P.

XX PR 18-NOV-2002; 2002US-0426742P.

XX PR 18-NOV-2002; 2002US-0427367P.

XX PA (AMHP) WYETH HOLDINGS CORP.

XX PI Murphy E, Projan SJ;

XX DR WPI; 2004-053616/05.

XX DR N-PSDB; ADJ27076.

XX PT Novel isolated Alloiooccus otitidis polypeptides useful for identifying

XX PT compounds that inhibit the activity or expression of the polypeptide and

XX PT thus are useful for treating infections caused by Alloiooccus otitidis.

XX PS Claim 7; SEQ ID NO 22; 433pp; English.

XX CC This invention relates to novel isolated Alloiooccus otitidis (A.

XX CC otitidis) nucleic acid molecules and encoded proteins thereof.

XX CC Specifically, it refers to proteins that are essential for the growth and

XX CC survival of the gram-positive bacterium A. otitidis, and hence provide

XX CC novel antibacterial targets. The present invention describes

XX CC pharmaceutical compositions and antisense compounds that are useful for

XX CC inhibiting activity or expression of these proteins. Furthermore, it

XX CC provides diagnostics and therapeutics that can be used to ameliorate

XX CC diseases that are associated with A. otitidis, such as otitis media with

XX CC effusion (OME) and various bacterial infections. Accordingly, these

XX CC compositions exhibit both antibacterial and antineurotic activities.

XX CC This polypeptide is an A. otitidis protein sequence of the invention.

XX SQ Sequence 400 AA;

Query Match 40.0%; Score 745; DB 8; Length 400;

Best Local Similarity 42.3%; Pred. No. 6e-57;

Matches 160; Conservative 68; Mismatches 146; Indels 4; Gaps 4;

QY 2 1QFSINRFLTHALNTTKRAISTKNAIPILSSIKIEVSTGVTLTGSNGQISIENTIPVS 61

DB 24 LKFSVKRTFELKVLKKVQIAVSSKSTIAITLKIEADNQGLTIGSNDSDISVESYLSVT 83

QY 62 NENAGLLITSPGAILLE-ASFFINIISSLPDISINVKIEHQHVLTSGKSEITLKGXDV 120

DB 84 DEGADLVIDEPGQIVLPQARLFANIVOKLPDTHFKVNVSSQGOQTQITSASASFTINGIDA 143

QY 121 DOYPLQEVSTENPLILTKLLKLSIAETAFAASLOESRPILTGVHIVLSNKHDPKAVAT 180

DB 144 MSYPLHPDIDLESFTPLPVDLFPKMINQTVIAVSNHESRPILTVGNLSLKEGR-LKAVAT 202

QY 181 DSHRMSQRLITLDNT-SADLMVVLPSKSLREPSAVFTDDIETVEFFPSQILFRSEHIS 239

DB 203 DSHRLSQRSIQLESAPDISFDIVIPGKSLTETLKMAD-EEVRVAISDNQILFALSQQ 261

QY 240 FYTRLEGNYPDTRLLMTEFEVFNQTSRSHAMERAFILSNATQNGTVKLEITQNH 299

DB 262 FYSRLLEGKYPTDRLIPGDTPTETITLDTKELQGVADRASLLSHEGKNVVLQTVTAEL 321

QY 300 SAHVNSPEVGKVNEDLDIVSSQGSDLTISFNPTYLIESLKAIKSETVKIHLFSPVRPTL 359


```

PD 29-JUL-1999.
PF 25-JAN-1999; 99WO-US001547.
XX
PR 27-JAN-1998; 98US-0074522P.
XX
PR 22-JUL-1998; 98US-0093727P.
XX
PA (UYRQ ) UNIV ROCKEFELLER.
XX
PI O'donnell ME, Zhang D, Whipple R;
XX
DR WPI; 1999-590685/50.
DR N-PSDB; AAZ31005.
XX
XX New isolated dnaE, dnaX and dnaB genes from Gram positive bacteria, used
PT to develop screening assays for identifying antibiotic compounds.
XX
PS Disclosure; Page 30-32; 132pp; English.
XX
CC This is the amino acid sequence encoded by the Staphylococcus aureus
CC partial dnaN gene. The invention relates to a number of isolated DNA
CC molecules from Gram positive bacterium, corresponding to dnaE (AAZ31001),
CC dnaX (AAZ31002), and dnaB (AAZ31003). The PolC, dnaN and dnaG genes
CC (AAZ31004-231006) are also identified. The dnaB gene corresponds to the
CC alpha subunit of the Escherichia coli, DNA polymerase III holoenzyme,
CC dnaX corresponds to the gamma and tau subunits, and dnaB corresponds to
CC the helicase. The alpha subunit is the actual DNA polymerase, the gamma
CC complex forms the clamp loader and tau is a "glue protein". dnaX encodes
CC both gamma and tau, Tau is the product of the full gene, while gamma is
CC the product of the first two thirds of the gene. dnaN forms the beta
CC subunit which forms the sliding clamp, and dnaG encodes a primase. The
CC DNA sequences of the invention can be used to identify agents that
CC inhibit or promote DNA replication by acting on various parts of the gram
CC positive bacterial DNA polymerase holoenzyme. The products and methods of
CC the invention can be used for identifying pharmacological agents or lead
CC compounds for agents active at the level of a replication protein
CC function, particularly DNA replication. The agents identified can be used
CC as antibiotics
XX
SQ Sequence 377 AA;
    Query Match          39.2%; Score 730.5; DB 2; Length 377;
    Best Local Similarity 39.3%; Pred. No. 1e-55;
    Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIOFSINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
Db 1 MMEFTIKRDYFITQLNDTLKALSPRTTLPILTGKIDAKEHEVILTSGDSEISITIPK 60
QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPISINVKIEHQHVLTSGKSEITLKGKD 119
Db 61 TVDGEDIVNISFTSGVLPGRFVDIIKKLPKGVKLTSTNEQFQTLITSGHSEFNLGLD 120
QY 120 VDQYPRLOVSTENPILKTLKLSIAETAFASLQESRPILTGTHVLSLHDKFKAVA 179
Db 121 PDQYLFPQVSDDDAIQSVKLVKRVIAQTNAFVSTETRPVLTGNVLQIEN-ELICTA 179
QY 180 TDSHRMSORLIPLDNTSADLMVLPKSLRPFSAVFTDDIETVEVFSPQLPSEHIS 239
Db 180 TDSHLAVRKQLQEDVSEKNVLIIPKALAEALNKIMSDNEEDIDIFFASNQVLFKGVNV 239
QY 240 FYTRLEGGYPTDRLLMTEFTEVVFNQSRHAMEAFILSNATONGTVKLEITQNH 299
Db 240 FISRLLEGHYPTDRLLMTEFTEVVFNQSRHAMEAFILSNATONGTVKLEITQNH 299
QY 300 SAHVNSPEVKYNEVDNDVSGSDLTISFNPTYLIESIKATKSETVKIHLSPVRPTL 359
Db 300 ELUSSTPEIGTVKEEVANDVEGSKISFNSKYMMDALKADINDEVEVEFGTMKPFIL 359
QY 360 TP-GDEERSFIQLITPVRT 377
Db 360 KPKGD--DSVTQLIIPRT 376

```

```

RESULT 22
AAU37133
ID AAU37133 standard; protein; 377 AA.
XX
AC AAU37133;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1303.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS54992.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 12726; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 377 AA;
    Query Match          39.2%; Score 730.5; DB 4; Length 377;
    Best Local Similarity 39.3%; Pred. No. 1e-55;
    Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIOFSINTLFIHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGQISIENTIPV 60
Db 1 MMEFTIKRDYFITQLNDTLKALSPRTTLPILTGKIDAKEHEVILTSGDSEISITIPK 60
QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPISINVKIEHQHVLTSGKSEITLKGKD 119

```

[illegible]

RESULT 23
AAU37107
ID AAU37107 standard; protein; 377 AA.

development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

RESULT 24
AAB31935
ID AAB31935 standard; protein; 377 AA.

PI O'donnell ME, Bruck I, Zhang D, Whipple R;
 XX WPI; 2001-147453/15.
 DR N-PSDB; AAF54735.
 XX
 XX Isolated DNA molecule from a Gram positive bacterium encoding DNA
 PT replication proteins used to identify compounds which have antibiotic
 PT activity.
 XX
 XX Disclosure; Page 40-41; 239pp; English.
 PS
 XX The present sequence represents a dnaN polypeptide. The specification
 CC describes DNA molecules from Gram positive bacteria, which comprise a
 CC coding region from a polC, dnaB, hoiA, hoiB, dnaX, ssb, dnaG or a
 CC dnaB gene. These sequences encode proteins that replicate the chromosome
 CC of Gram positive bacteria. They are used for sequencing and amplification
 CC of DNA and in drug discovery to identify compounds which have antibiotic
 CC activity through interference with replication. They are used in methods
 CC for identifying compounds that are active at the level of DNA replication
 CC and result in arrest of cell growth or cell death of bacteria to treat
 CC bacterial infections in animals
 XX
 SQ Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 4; Length 377;
 Best Local Similarity 39.3%; Pred. No. 1e-55;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIOFSINRTFLHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGOISIENTIPV 60
 DB 1 MMEFTIKRDYFITQNDTLKAIISPTTLITGKIDAKEHEVILTGSDSEISIEITPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPDISINVKIEIQHQVVLTSKSEITLKGKD 119
 DB 61 TVDGEDIVNISGTVLPGRFFVDIIKKPGKVKLSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDQYPRQEVSTENPLIKTKLLKSIIAETAPASLQESRPILITGVHIVLSNKHDFKAVA 179
 DB 121 PDQYPLLPQVSRDDAIQLSVKLVKNVIAQTNFVSTSTTRPVLITGVNWLIOEN-ELICTA 179
 QY 180 TDSHRMSORLITLONTSDADLMVLPKSLRPSFAVFTDDIEVVEFFSPQILFRSEHIS 239
 DB 180 TDSHRLAVRKQLQEDVSENKNVIIIPGKALAEINKIMSDNEEDIDIFFASNQVLFKGVN 239
 QY 240 FYTRLLEGNYPDTRLMLTEFETEVVFNTOQLSRHAMERAFILSNATONGTVKLEITQNH 299
 DB 240 FISRLLEGHYPDTRLFPENVEIKLSIDNGEYFHAIDRASLLAREGGNNVILKSTGDDV 299
 QY 300 SAHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSTVKIHFSLPVRPFTL 359
 DB 300 ELSSTSPBIGTVKEEVDANDVEGGLSKISFNKYMMDALKADINDNEVEVEFFGTMKPFIL 359
 QY 360 TP-GDEESFQILTPVRT 377
 DB 360 KPKG-D-DSVTQLILPIRT 376

RESULT 25
 ABB81847
 ID ABB81847 standard; protein; 377 AA.
 XX
 XX ABB81847;
 AC
 XX
 XX 20-SEP-2002 (first entry)
 DT
 XX Staphylococcus aureus STAAU_R2.
 DE
 XX STAAU-R2; drug screening; antibacterial; bacterial infection.
 KW
 XX Staphylococcus aureus.
 OS
 XX WO200244718-A2.
 FN
 XX

PD 06-JUN-2002.
 XX 30-NOV-2001; 2001WO-CA001754.
 XX
 XX 30-NOV-2000; 2000US-00727892.
 PR 20-JUN-2001; 2001US-00885561.
 XX
 XX (PHAG-) PHAGETECH INC.
 PA
 XX Pelletier J, Gros P, Dubow M;
 FI
 XX WPI; 2002-566592/60.
 DR N-PSDB; ABN85787.
 XX
 XX Use of a Staphylococcus aureus (STAAUR2) polypeptide or composition
 CC comprising STAAUR2 and polypeptide encoded by bacteriophage open reading
 CC frame that specifically interacts with STAAUR2, for identifying a
 CC compound active on STAAUR2.
 PT
 XX Claim 1; Fig 1; 131pp; English.
 PS
 XX The invention relates to a novel Staphylococcus aureus polypeptide
 CC (STAAU_R2), and it's use in drug screening assays. The polypeptide of the
 CC invention has antibacterial activity. The compound active on STAAU_R2 is
 CC useful for the manufacture of an antibacterial agent or for the
 CC manufacture for medication for treating or preventing a bacterial
 CC infection. The sequence represents the S. aureus STAAU_R2 of the
 CC invention
 XX
 SQ Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 5; Length 377;
 Best Local Similarity 39.3%; Pred. No. 1e-55;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 QY 1 MIOFSINRTFLHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTSGNGOISIENTIPV 60
 DB 1 MMEFTIKRDYFITQNDTLKAIISPTTLITGKIDAKEHEVILTGSDSEISIEITPK 60
 QY 61 SNEAGLL-ITSPGAILLEASFFINISSLDPDISINVKIEIQHQVVLTSKSEITLKGKD 119
 DB 61 TVDGEDIVNISGTVLPGRFFVDIIKKPGKVKLSTNEQFQTLITSGHSEFNLSGLD 120
 QY 120 VDQYPRQEVSTENPLIKTKLLKSIIAETAPASLQESRPILITGVHIVLSNKHDFKAVA 179
 DB 121 PDQYPLLPQVSRDDAIQLSVKLVKNVIAQTNFVSTSTTRPVLITGVNWLIOEN-ELICTA 179
 QY 180 TDSHRMSORLITLONTSDADLMVLPKSLRPSFAVFTDDIEVVEFFSPQILFRSEHIS 239
 DB 180 TDSHRLAVRKQLQEDVSENKNVIIIPGKALAEINKIMSDNEEDIDIFFASNQVLFKGVN 239
 QY 240 FYTRLLEGNYPDTRLMLTEFETEVVFNTOQLSRHAMERAFILSNATONGTVKLEITQNH 299
 DB 240 FISRLLEGHYPDTRLFPENVEIKLSIDNGEYFHAIDRASLLAREGGNNVILKSTGDDV 299
 QY 300 SAHVNSPEVKVKNEDLDIVSQSGDLTISFNPTYLIESLKAIKSTVKIHFSLPVRPFTL 359
 DB 300 ELSSTSPBIGTVKEEVDANDVEGGLSKISFNKYMMDALKADINDNEVEVEFFGTMKPFIL 359
 QY 360 TP-GDEESFQILTPVRT 377
 DB 360 KPKG-D-DSVTQLILPIRT 376

RESULT 26
 ABB16160
 ID ABB16160 standard; protein; 377 AA.
 XX
 XX ABB16160;
 AC
 XX
 XX 19-JUN-2003 (first entry)
 DT
 XX Protein encoded by Prokaryotic essential gene #1687.
 DE

XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW	
XX	
OS	Staphylococcus aureus.
XX	
PN	WO200277183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
XX	(ELIT-) ELITRA PHARM INC.
PA	
XX	
PI	Wang L, Zamudio C, Malone C, Zyskind JW;
PI	Wall D, Trawick JD, Carr GJ, Xu HH;
XX	Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
DR	N-PSDB; ACA20030.
DR	

RESULT	27
ABM72805	
ID	ABM72805 standard; protein; 377 AA.
XX	
XX	ABM72805;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Staphylococcus aureus protein #2045.
XX	
KW	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW	enzymatic assay; antibiotic target.
XX	
OS	Staphylococcus aureus.

XX	ABM72805;	
XX	AC	
XX	XX	
XX	XX	
DT	20-NOV-2003	(first entry)
XX	XX	
XX	XX	
DE	Staphylococcus aureus	protein #2045.
XX	XX	
XX	Antibacterial;	vaccine; gene therapy; infection; sepsis; diagnosis;
KW	enzymatic assay;	antibiotic target.
XX	XX	
OS	Staphylococcus aureus.	
XX	XX	
XX	XX	
PN	WO200294868-A2.	
XX	XX	
XX	XX	
PD	28-NOV-2002.	

XX	WO200294868-A2.
PN	
XX	
PD	28-NOV-2002.
XX	
XX	27-MAR-2002; 2002WO-IB002637.
PF	
XX	
XX	27-MAR-2001; 2001GB-00007661.
PR	
XX	(CHIR-) CHIRON SPA.
PA	
XX	
XX	Masignani V, Mora M, Scarselli M;
PI	
XX	
XX	WPI; 2003-120786/11.
DR	N-PSDB; ACF74365.
XX	
PT	New Staphylococcus aureus protein, useful as a vaccine for treating or
PT	preventing Staphylococcal infection, specifically an infection caused by
PT	S. aureus, e.g. sepsis.
XX	
PS	Claim 1; SEQ ID NO 4090; 49pp; English.

CC Staphylococcus aureus. A composition comprising the *S. aureus* protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to *Staphylococcus* bacteria, specifically an
CC infection caused by *S. aureus*. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel *S. aureus* proteins of the invention
XX
XX Sequence 377 AA:
SQ

useful as a pharmaceutical, particularly as a vaccine for treating or preventing infection due to *Staphylococcus* bacteria, specifically an infection caused by *S. aureus*. The composition is particularly useful for treating or preventing sepsis in a patient. The composition can also be used for diagnostics. The protein is also used in an assay for enzymatic studies and as a target for antibiotics. This sequence represents one of the novel *S. aureus* proteins of the invention

CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel *S. aureus* proteins of the invention
XX
SQ Sequence 377 AA:

Query Match 39.2%; Score 730.5; DB 6; Length 377;
Best Local Similarity 39.3%; Pred. No. 1e-55;
Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIQFSNRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSVGVLTTSGNGOISIENTIPV 60
DB 1 MMEFTIKRDYFITQLNDTLKAISPRTTLPITLGKIDAKEHEVLTGSDSEISIEIPIPK 60
QY 61 SNENAGLL-ITSFGAILLEASFFNIISLDPDISINVKIEIQHQVLTSGKSEITLKGD 119
DB 61 TVDGEDIVNISGTGSVWLPGRFFVDIIKKLPKDKVCLSTNEQFOTLITSGHSEFNLSGLD 120
QY 120 VDOYPRLOEVSTENPLILKTLLKSIIAETAFASLQESRPILTGVIHLVSNHKDFKAVA 179
DB 121 PDQYPLLQVSRDDAIQISVKLVKNVIAQTFAVSTSETRPLVTGVNWLIOEN-ELICTA 179
QY 180 TDSHRMSORLIITLDNTSADLMVWLPSKSLREFSFAVFTDDIETVEVFSPSQILFRSHIS 239
DB 180 TDHRLAVRKUQLEDVSENKVIIIPGKALAELNKIMSDNEEDIIDIFFASNQVLPKGVNVN 239
QY 240 FYTRLLEGNYPTDRLLMTFEFETEVEVFNTQSRLHAMERAFILSNATONGTVKLEITQNHI 299
DB 240 FISRLEGHYPDTRLFENYEIKLSINDGEFYHAIDRASLAREGGNVNIKLSGTGDVV 299
QY 300 SAHVNSPEVGKVNEDLDIVSOGSDLTISFNPTYLIESLKAISKETWKIHFLSPVRPFTL 359
DB 300 ELSSTSPEIGTGVKEEVDANDVEGSLGISFNSKYMMALKAINDEVEVEFFGTMKPFI 359
QY 360 TP-GDEESFIQILTPVRT 377
DB 360 KPKG--DSVTQLILPIRT 376

RESULT 28
ADD26272
ID ADD26272 standard; protein; 377 AA.
AC
XX ADD26272;
DT 15-JAN-2004 (first entry)
DX Staphylococcus aureus dnaN protein SEQ ID NO:77.
DE
XX crystallised recombinant polypeptide; Staphylococcus aureus;
KW Streptococcus pneumoniae; Helicobacter pylori; Pseudomonas aeruginosa;
KW nucleic acid processing; crystal form;
KW ribonuclease diphosphate reductase; nrde;
KW polyribonucleotide phosphorylase; pnpA;
KW transcription termination factor Rho;
KW putative polynucleotide nucleotidyltransferase; DNA ligase;
KW DNA polymerase III; dnaN;
KW ribonucleoside diphosphate reductase minor subunit; uridylylate kinase;
KW orotate phosphoribosyltransferase; pyrE; uracil-DNA-glycosylase;
KW antibacterial; vaccine.

Staphylococcus aureus..
OS
XX
OS
XX
PN W02003025004-A2.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-CA001411.
XX
PR 18-SEP-2001; 2001US-0323040P.
PR 27-SEP-2001; 2001US-0325307P.
PR 27-SEP-2001; 2001US-0325421P.
PR 28-SEP-2001; 2001US-0325891P.
PR 01-OCT-2001; 2001US-0326337P.
PR 03-OCT-2001; 2001US-0326774P.
PR 04-OCT-2001; 2001US-0327193P.
PR 30-OCT-2001; 2001US-0340922P.
PR 05-NOV-2001; 2001US-0338709P.

PR 06-NOV-2001; 2001US-0333269P.
PR 18-DEC-2001; 2001US-0341679P.
PA (AFFI-) AFFINIUM PHARM INC.
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowemith C, Awrey D;
PI Beattie B, Charandien V, Cox B, Domagala M, Houston S, Li Q;
PI Nethery K, Ng I, Ouyang H, Pinder B, Sheldrick B, Viola C;
PI Wrezel O;
XX WPI: 2003-371793/35.
DR N-PSDB; ADD26271.
XX
XX New crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae, Helicobacter pylori, Pseudomonas aeruginosa
PT involved in nucleic acid processing, useful as targets for pathogenic
PT bacteria.
XX Claim 218; SEQ ID NO 77; 298pp; English.
XX The present invention describes a crystallised recombinant polypeptide
CC (I) comprising the amino acid sequence of polypeptides from
CC Staphylococcus aureus, Streptococcus pneumoniae, Helicobacter pylori and
CC Pseudomonas aeruginosa which are involved in nucleic acid processing, or
CC amino acid sequences having at least 90 % identity with the polypeptide
CC sequence, where the polypeptide is in crystal form. (I) comprises the
CC amino acid sequence (S) of polypeptides involved in nucleic acid
CC processing, which includes ribonuclease diphosphate reductase, major
CC subunit (nrde), polyribonucleotide phosphorylase (pnpA), transcription
CC termination factor Rho (Rho), putative polynucleotide
CC nucleotidyltransferase (pnp), DNA ligase (lig), DNA polymerase III, beta-
CC subunit (dnaN) or ribonucleoside diphosphate reductase minor subunit (R2)
CC (nrdf) from S. aureus; uridylylate kinase (pyrE) or uracil-DNA-glycosylase
CC (lung) from P. aeruginosa; and nrde from S. pneumoniae. (I) is in a
CC crystal form. (I) has antibacterial activity, and can be used in
CC vaccines. (I) is useful for designing a modulator for the prevention or
CC treatment of S. aureus, S. pneumoniae, H. pylori and P. aeruginosa
CC related diseases or disorders. The method comprises providing a three-
CC dimensional (3D) structure for (I), identifying a potential modulator by
CC reference to the 3D structure, contacting the potentially modulator with
CC the recombinant polypeptide and assaying the activity of the polypeptide
CC or determining the viability of S. aureus, S. pneumoniae, H. pylori or P.
CC aeruginosa after contact with the modulator, where a change in the
CC activity of the polypeptide or the viability of the bacteria indicates
CC that the modulator may be useful for preventing or treating the disease
CC or disorder. The structural and functional information of (I) aid in the
CC discovery and design of therapeutic and diagnostic molecules. The present
CC discovery is used in the exemplification of the present invention.
XX
SQ -Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 7; Length 377;
Best Local Similarity 39.3%; Pred. No. 1e-55;
Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIQFSNRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSVGVLTTSGNGOISIENTIPV 60
DB 1 MMEFTIKRDYFITQLNDTLKAISPRTTLPITLGKIDAKEHEVLTGSDSEISIEIPIPK 60
QY 61 SNENAGLL-ITSFGAILLEASFFNIISLDPDISINVKIEIQHQVLTSGKSEITLKGD 119
DB 61 TVDGEDIVNISGTGSVWLPGRFFVDIIKKLPKDKVCLSTNEQFOTLITSGHSEFNLSGLD 120
QY 120 VDOYPRLOEVSTENPLILKTLLKSIIAETAFASLQESRPILTGVIHLVSNHKDFKAVA 179
DB 121 PDQYPLLQVSRDDAIQISVKLVKNVIAQTFAVSTSETRPLVTGVNWLIOEN-ELICTA 179
QY 180 TDSHRMSORLIITLDNTSADLMVWLPSKSLREFSFAVFTDDIETVEVFSPSQILFRSHIS 239
DB 180 TDHRLAVRKUQLEDVSENKVIIIPGKALAELNKIMSDNEEDIIDIFFASNQVLPKGVNVN 239
QY 240 FYTRLLEGNYPTDRLLMTFEFETEVEVFNTQSRLHAMERAFILSNATONGTVKLEITQNHI 299

Query Match 39.2%; Score 730.5; DB 6; Length 377;
Best Local Similarity 39.3%; Pred. No. 1e-55;
Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIQFSNRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSVGVLITSGNGOISIENTIPV 60
DB 1 MMEFTIKRDYFITQLNDTLKAISPRTTLPILTIGIKIDAKEHEVLTGSDSEISIEIPIPK 60
QY 61 SNENAGLL-ITSFGAILLEASFFNIISLDPDISINVKIEIQHQVLTSGKSEITLKGD 119
DB 61 TVDGEDIVNISGTGSVWLPGRFFVDIIKKLPKGVKLSTNEQFOTLITSGHSEFNLSGLD 120
QY 120 VDOYPRLOEVSTENPLILKTLLKSIIAETAFASLQESRPILTGVIHVLNSHKDFKAVA 179
DB 121 PDQYPLLQVSRDDAIQISVKLVKNVIAQTFAVSTSETRPLVTGVNWLIOEN-ELICTA 179
QY 180 TDSHRMSORLIITLDNTSADLMVWLPSKSLREFSFAVFTDDIETVEVFSPSQILFRSHIS 239
DB 180 TDHRLAVRKUQLEDVSENKVIIIPGKALAELNKIMSDNEEDIIDIFFASNQVLPKGVNVN 239
QY 240 FYTRLLEGNYPTDRLMLTFEFTEVFVNFTQSRLHAMERAFILSNATONGTVKLEITQNHI 299
DB 240 FISRLEGHYPDTRLFENYEIKLSINDGEFYHAIDRASLAREGGNVNIKLGSTDVV 299
QY 300 SAHVNSPEVGKYNEDLDIVSOGSDLTISFNPTYLIESLKAISKETWKIHFLSPVRPFTL 359
DB 300 ELSSTSPEIGTGVKEEVDANDVEGSLGISFNSKYMMALKAINDEVEVEFFGTMKPFI 359
QY 360 TP-GDEESFIQILTPVRT 377
DB 360 KPKG--DSVTQLILPIRT 376

RESULT 28
ADD26272
ID ADD26272 standard; protein; 377 AA.
AC
XX ADD26272;
DT 15-JAN-2004 (first entry)
DX Staphylococcus aureus dnaN protein SEQ ID NO:77.
DE
XX crystallised recombinant polypeptide; Staphylococcus aureus;
KW Streptococcus pneumoniae; Helicobacter pylori; Pseudomonas aeruginosa;
KW nucleic acid processing; crystal form;
KW ribonuclease diphosphate reductase; nrde;
KW polyribonucleotide phosphorylase; pnpA;
KW transcription termination factor Rho;
KW putative polynucleotide nucleotidyltransferase; DNA ligase;
KW DNA polymerase III; dnaN;
KW ribonucleoside diphosphate reductase minor subunit; uridylylate kinase;
KW orotate phosphoribosyltransferase; pyrE; uracil-DNA-glycosylase;
KW antibacterial; vaccine.

Staphylococcus aureus..
OS
XX
OS
XX
PN W02003025004-A2.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-CA001411.
XX
XX 18-SEP-2001; 2001US-0323040P.
PR 27-SEP-2001; 2001US-0325307P.
PR 27-SEP-2001; 2001US-0325421P.
PR 28-SEP-2001; 2001US-0325891P.
PR 01-OCT-2001; 2001US-0326337P.
PR 03-OCT-2001; 2001US-0326774P.
PR 04-OCT-2001; 2001US-0327193P.
PR 30-OCT-2001; 2001US-0340922P.
PR 05-NOV-2001; 2001US-0338709P.

PR 06-NOV-2001; 2001US-0333269P.
PR 18-DEC-2001; 2001US-0341679P.
PA (AFFI-) AFFINIUM PHARM INC.
XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowemith C, Awrey D;
PI Beattie B, Charandien V, Cox B, Domagala M, Houston S, Li Q;
PI Nethery K, Ng I, Ouyang H, Pinder B, Sheldrick B, Viola C;
PI Wrezel O;
XX WPI: 2003-371793/35.
DR N-PSDB; ADD26271.
XX
XX New crystallized recombinant polypeptides from Staphylococcus aureus,
PT Streptococcus pneumoniae, Helicobacter pylori, Pseudomonas aeruginosa
PT involved in nucleic acid processing, useful as targets for pathogenic
PT bacteria.
XX Claim 218; SEQ ID NO 77; 298pp; English.
XX The present invention describes a crystallised recombinant polypeptide
CC (I) comprising the amino acid sequence of polypeptides from
CC Staphylococcus aureus, Streptococcus pneumoniae, Helicobacter pylori and
CC Pseudomonas aeruginosa which are involved in nucleic acid processing, or
CC amino acid sequences having at least 90 % identity with the polypeptide
CC sequence, where the polypeptide is in crystal form. (I) comprises the
CC amino acid sequence (S) of polypeptides involved in nucleic acid
CC processing, which includes ribonuclease diphosphate reductase, major
CC subunit (nrde), polyribonucleotide phosphorylase (pnpA), transcription
CC termination factor Rho (Rho), putative polynucleotide
CC nucleotidyltransferase (pnp), DNA ligase (lig), DNA polymerase III, beta-
CC subunit (dnaN) or ribonucleoside diphosphate reductase minor subunit (R2)
CC (nrdf) from S. aureus; uridylylate kinase (pyrE) or orotate-
CC phosphoribosyltransferase (pyrE) from H. pylori; uracil-DNA-glycosylase
CC (lung) from P. aeruginosa; and nrde from S. pneumoniae. (I) is in a
CC crystal form. (I) has antibacterial activity, and can be used in
CC vaccines. (I) is useful for designing a modulator for the prevention or
CC treatment of S. aureus, S. pneumoniae, H. pylori and P. aeruginosa
CC related diseases or disorders. The method comprises providing a three-
CC dimensional (3D) structure for (I), identifying a potential modulator by
CC reference to the 3D structure, contacting the potentially modulator with
CC the recombinant polypeptide and assaying the activity of the polypeptide
CC or determining the viability of S. aureus, S. pneumoniae, H. pylori or P.
CC aeruginosa after contact with the modulator, where a change in the
CC activity of the polypeptide or the viability of the bacteria indicates
CC that the modulator may be useful for preventing or treating the disease
CC or disorder. The structural and functional information of (I) aid in the
CC discovery and design of therapeutic and diagnostic molecules. The present
CC discovery is used in the exemplification of the present invention.
XX
SQ -Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 7; Length 377;
Best Local Similarity 39.3%; Pred. No. 1e-55;
Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

QY 1 MIQFSNRTLFTHALNTTKRAISTKNAIPILSSIKIEVTSVGVLITSGNGOISIENTIPV 60
DB 1 MMEFTIKRDYFITQLNDTLKAISPRTTLPILTIGIKIDAKEHEVLTGSDSEISIEIPIPK 60
QY 61 SNENAGLL-ITSFGAILLEASFFNIISLDPDISINVKIEIQHQVLTSGKSEITLKGD 119
DB 61 TVDGEDIVNISGTGSVWLPGRFFVDIIKKLPKGVKLSTNEQFOTLITSGHSEFNLSGLD 120
QY 120 VDOYPRLOEVSTENPLILKTLLKSIIAETAFASLQESRPILTGVIHVLNSHKDFKAVA 179
DB 121 PDQYPLLQVSRDDAIQISVKLVKNVIAQTFAVSTSETRPLVTGVNWLIOEN-ELICTA 179
QY 180 TDSHRMSORLIITLDNTSADLMVWLPSKSLREFSFAVFTDDIETVEVFSPSQILFRSHIS 239
DB 180 TDHRLAVRKUQLEDVSENKVIIIPGKALAELNKIMSDNEEDIIDIFFASNQVLPKGVNVN 239
QY 240 FYTRLLEGNYPTDRLMLTFEFTEVFVNFTQSRLHAMERAFILSNATONGTVKLEITQNHI 299

Db 240 FISRLLEGHPDTRFLFPENYEIKLSIDNGEFVHAIDRASLLAREGNNVILKSTGDDVV 299
 Qy 300 SAHVNSEVGVKNEEDLIVSQSGLTISNPNYLSKAISKSETVKHFLSPVPPFTL 359
 Db 300 ELSTSPEIGTVKEEVANDVEGSLKISFNSKYMMDALKADNDDEVEVEFFGTMKPFIL 359
 Qy 360 TP-GDEESFIQLITPVRT 377
 Db 360 KPKGDD--DSVTQLILPRT 376
 ID ADD26274
 XX 240 FISRLLEGHPDTRFLFPENYEIKLSIDNGEFVHAIDRASLLAREGNNVILKSTGDDVV 299
 AC ADD26274;
 DT 15-JAN-2004 (first entry)
 XX Staphylococcus aureus dnaN protein SEQ ID NO:79.
 DE crystallized recombinant polypeptide; Staphylococcus aureus;
 KW Streptococcus pneumoniae; Helicobacter pylori; Pseudomonas aeruginosa;
 KW nucleic acid processing; crystal form;
 KW ribonuclease diphosphate reductase; nrdE;
 KW polyribonucleotide phosphorylase; pnpA;
 KW transcription termination factor Rho;
 KW putative polynucleotide nucleotidyltransferase; DNA ligase;
 KW DNA polymerase III; dnaN;
 KW ribonucleoside diphosphate reductase minor subunit; uridylylate kinase;
 KW orotate phosphoribosyltransferase; pyrE; uracil-DNA-glycosylase;
 KW antibacterial; vaccine.
 XX Staphylococcus aureus.

RESULT 29

ID ADD26274
 XX 240 FISRLLEGHPDTRFLFPENYEIKLSIDNGEFVHAIDRASLLAREGNNVILKSTGDDVV 299
 AC ADD26274;
 DT 15-JAN-2004 (first entry)
 XX Staphylococcus aureus dnaN protein SEQ ID NO:79.
 DE crystallized recombinant polypeptide; Staphylococcus aureus;
 KW Streptococcus pneumoniae; Helicobacter pylori; Pseudomonas aeruginosa;
 KW nucleic acid processing; crystal form;
 KW ribonuclease diphosphate reductase; nrdE;
 KW polyribonucleotide phosphorylase; pnpA;
 KW transcription termination factor Rho;
 KW putative polynucleotide nucleotidyltransferase; DNA ligase;
 KW DNA polymerase III; dnaN;
 KW ribonucleoside diphosphate reductase minor subunit; uridylylate kinase;
 KW orotate phosphoribosyltransferase; pyrE; uracil-DNA-glycosylase;
 KW antibacterial; vaccine.
 XX Staphylococcus aureus.

OS WO2003025004-A2.
 XX 27-MAR-2003.
 XX 18-SEP-2002; 2002WO-CA001411.
 PR 18-SEP-2001; 2001US-0323040P.
 PR 27-SEP-2001; 2001US-032307P.
 PR 27-SEP-2001; 2001US-0325421P.
 PR 28-SEP-2001; 2001US-0325891P.
 PR 01-OCT-2001; 2001US-0326337P.
 PR 04-OCT-2001; 2001US-0326774P.
 PR 30-OCT-2001; 2001US-0327193P.
 PR 05-NOV-2001; 2001US-0340922P.
 PR 06-NOV-2001; 2001US-0338709P.
 PR 18-DEC-2001; 2001US-0341679P.
 XX (AFFI-) AFFINIUM PHARM INC.
 XX Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowsmith C, Awrey D;
 PI Beattie B, Canadian V, Cox B, Domagala M, Houston S, Li Q;
 PI Netherly K, Ng I, Ouyang H, Pinder B, Sheldrick B, Viola C;
 PI Wrezel O;
 XX WPI; 2003-371793/35.
 DR N-PSDB; ADD26273.
 XX New crystallized recombinant polypeptides from Staphylococcus aureus,
 PT Streptococcus pneumoniae, Helicobacter pylori, Pseudomonas aeruginosa
 PT involved in nucleic acid processing, useful as targets for pathogenic
 PT bacteria.
 XX Claim 218; SEQ ID NO 79; 298pp; English.
 PS The present invention describes a crystallised recombinant polypeptide
 CC (I) comprising the amino acid sequence of polypeptides from

CC Staphylococcus aureus, Streptococcus pneumoniae, Helicobacter pylori and
 CC Pseudomonas aeruginosa which are involved in nucleic acid processing, or
 CC amino acid sequences having at least 90 % identity with the polypeptide
 CC sequence, where the polypeptide is in crystal form. (I) comprises the
 CC amino acid sequence (S) of polypeptides involved in nucleic acid
 CC processing, which includes ribonuclease diphosphate reductase, major
 CC subunit (nrdE), polyribonucleotide phosphorylase (pnpA), transcription
 CC termination factor Rho (Rho), putative polynucleotide
 CC nucleotidyltransferase (pnp), DNA ligase (lig), DNA polymerase III, beta-
 CC subunit (dnaN) or ribonucleoside diphosphate reductase minor subunit (R2)
 CC (nrdF) from S. aureus; uridylylate kinase (pyrE) or orotate
 CC phosphoribosyltransferase (pyrE) from H. pylori; uracil-DNA-glycosylase
 CC (ung) from P. aeruginosa; and nrdE from S. pneumoniae. (I) is in a
 CC crystal form. (I) has antibacterial activity, and can be used in a
 CC vaccines. (I) is useful for designing a modulator for the prevention or
 CC treatment of S. aureus, S. pneumoniae, H. pylori and P. aeruginosa
 CC related diseases or disorders. The method comprises providing a three-
 CC dimensional (3D) structure for (I), identifying a potential modulator by
 CC reference to the 3D structure, contacting the potential modulator with
 CC the recombinant polypeptide and assaying the activity of the polypeptide
 CC or determining the viability of S. aureus, S. pneumoniae, H. pylori or P.
 CC aeruginosa after contact with the modulator, where a change in the
 CC activity of the polypeptide or the viability of the bacterium indicates
 CC that the modulator may be useful for preventing or treating the disease
 CC or disorder. The structural and functional information of (I) aid in the
 CC discovery and design of therapeutic and diagnostic molecules. The present
 CC sequence is used in the exemplification of the present invention.
 XX

Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 7; Length 377;
 Best Local Similarity 39.3%; Pred. No. 1e-55;
 Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;
 Qy 1 MIQPSINRTLFHAIKNTKKAISTKNAIPILSSIKIEVSTGVLTSNGQISIENTIPV 60
 Db 1 MMEFTIKRDYFITQLNDTLKAIPTPTLTILTKIDAKEHEVLTGSDSEISIEITPK 60
 Qy 61 SNENAGLL-ITSPGAILLEAFSEFINIISLPDISINVEIEHQVLTSGKSEITLKGKD 119
 Db 61 TVGEDIWISGTVGVLPGRFVDIIKLPKDVKLSTNEQFTLTITSGHSEFNLSGLD 120
 Qy 120 VDQYPRLEQVSTENPLIKTKLKSIIAETAFAASLQESRPILTGVHIVLSNKHDKFAVA 179
 Db 121 PDQYPLLPQVSRDDAIQLSVKLVKNVIAQINFAVSTSETPVLTGVNWLQEN-ELICTA 179
 Qy 180 TDSHRKSQRLITLNTSADLMVLPKSLREFSAVTDITVEVTFSPSQILFRSEHIS 239
 Db 180 TDSHRLAVRKLQLEDVSENKNVPIPGKALAEKINSNEEDIDIFFASNQVLFKVGNVN 239
 Qy 240 FYTRLLEGNYPDTRLLMTFETEVVFNQSLRHAMERAFILSNATQNGTVKLEITONHI 299
 Db 240 FISRLLEGHPDTRFLFPENYEIKLSIDNGEFVHAIDRASLLAREGNNVILKSTGDDVV 299
 Qy 300 SAHVNSEVGVKNEEDLIVSQSGLDITISNPNYLSKAISKSETVKHFLSPVPPFTL 359
 Db 300 ELSTSPEIGTVKEEVANDVEGSLKISFNSKYMMDALKADNDDEVEVEFFGTMKPFIL 359
 Qy 360 TP-GDEESFIQLITPVRT 377
 Db 360 KPKGDD--DSVTQLILPRT 376

RESULT 30

ABW01648
 ID ABW01648 standard; protein; 377 AA.
 XX AC ABW01648;
 XX 12-FEB-2004 (first entry)
 XX Staphylococcus aureus dnaN gene beta subunit protein.

KW Polymerase III enzyme; dnaE; dnaX; dnaB; Gram positive bacteria;
 XX drug discovery; antibiotic activity.

OS Staphylococcus aureus.

XX US2003129633-A1.

XX 10-JUL-2003.

XX 28-OCT-2002; 2002US-00282287.

XX 13-FEB-1998; 98US-0074572P.

XX 22-JUL-1998; 98US-0093727P.

XX 22-JAN-1999; 99US-00235245.

XX (ODON/) O'DONNELL M E.

XX (ZHAN/) ZHANG D.

XX (WHIP/) WHIPPLE R.

XX O'donnell ME, Zhang D, Whipple R;

XX WPI; 2003-829557/77.

XX N-PSDB; AAD62920.

XX New DNA replication proteins (i.e. subunits of the Staphylococcus aureus

XX DNA polymerase III enzyme) and genes, useful in drug discovery to screen
 XX large libraries of chemicals for identification of compounds with
 XX antibiotic activity.
 XX Disclosure; Page 18-19; 69pp; English.

XX The invention relates to an isolated polypeptide, which comprises at
 XX least one functionally active subunit of a Staphylococcus aureus DNA
 XX polymerase III enzyme. The subunit comprises a 573 residue dnaE amino
 XX acid sequence, a 566 residue dnaX amino acid sequence and/or a 457
 XX residue dnaB amino acid sequence. The proteins and nucleic acids
 XX replicate the chromosome of Gram positive bacteria and are useful in drug
 XX discovery to screen large libraries of chemicals for identification of
 XX compounds with antibiotic activity. The present sequence is S. aureus
 XX dnaN gene beta subunit protein

XX Sequence 377 AA;

Query Match 39.2%; Score 730.5; DB 7; Length 377;

Best Local Similarity 39.3%; Pred. No. 1e-55;

Matches 149; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

Qy 1 MIQSFINTLETHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPV 60

Db 1 MMEFTIKRDYFITQLNDTLKAISPRTPITLITGKIDAKEHEVILTSGDSSEISIEITPK 60

Qy 61 SNEAGLL-ITSPGAILLEAGFFINISLDPDISINVKIEHQHVLTSGKSEITLKGKD 119

Db 61 TVDGEDVINISGTVSVLPGRFFVDIIKLPKQDVLTSTNQFQTLITSGHSEFNLGLD 120

Qy 120 VDQYPRLOEVSTENPLILTKLLKLSIIAETAFAASLOESRPILTGVHVLNSHDKFKAVA 179

Db 121 PDQYPLLQVSRDADAQLSVKLVKNVIAQTNPFAVSTSETRPVLTVGNWLIQEN-ELICTA 179

Qy 180 TDSHRMSORLITLNTSADLMVLPSKLSIRFSAVFTDDETVEVFPSQLFRSEHIS 239

Db 180 TDSHRLAVRKLEQVSENVKNVIFPKALAEINKIMSDNEEDIDIFFASNOVLFRKGVNV 239

Qy 240 FYTRLLEGNYPDTRLLETEPETEVFNTQSLRHAMERAFILSNATQNGTVKLEITQNH 299

Db 240 FISRLLEGHPDTRLLEFENYEIKLSIDNGEFYHAIDRASLLAREGGNNVILKSGDDVV 299

Qy 300 SAHVNSPVGKVNEDLDIVSQSGDLTISFNPTYLIESLKAISKSTVKIHLSPVRPFTL 359

Db 300 ELSSTSPBIGTVEVDANDVEGSLKISFNSKYMDALKAINDEVEVEFFGTMKPFIL 359

Qy 360 TP-GDEESFQTLITPVRT 377

||| : ||| ||| |||

Db 360 KPKGDD--DSVTQLILPIRT 376

XX RESULT 31

XX AAU34173

XX ID AAU34173 standard; protein; 376 AA.

XX AC AAU34173;

XX DT 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #449.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 XX antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 XX N-PSDB; AAS52032.

XX New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 5669; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 XX prokaryotic cellular proliferation, their use in identifying the genes,
 XX their use in the discovery of novel antibiotics, the essential genes,
 XX themselves and the encoded proteins. The prokaryotes used are Escherichia
 XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 XX useful for the identification of potential new targets for antibiotic
 XX development. The antisense nucleic acids can also be used to identify
 XX proteins used in proliferation, to express these proteins, and to obtain
 XX antibodies capable of binding to the expressed proteins. The proteins can
 XX be used to screen compounds in rational drug discovery programmes. The
 XX antisense nucleic acid sequence is also useful to screen for homologous
 XX nucleic acids which are required for cell proliferation in a wide variety
 XX of organisms. The present sequence represents an essential prokaryotic
 XX cellular proliferation protein. Note: The sequence data for this patent
 XX did not form part of the printed specification, but was obtained in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 376 AA;

Query Match 38.9%; Score 725.5; DB 4; Length 376;

Best Local Similarity 39.2%; Pred. No. 2.9e-55;

Matches 148; Conservative 90; Mismatches 135; Indels 5; Gaps 4;

Qy 2 IQSFINTLETHALNTTKRAISTKNAIPILSSIKIEVTSTGVTLTGSGNGQISIENTIPVS 61

Db 1 MEFTIKRDYFITQLNDTLKAISPRTPITLITGKIDAKEHEVILTSGDSSEISIEITPK 60

PI	Buchrieser C, Frangeul L, Couve E, Rundiok C, Psihi H, Dehoux P;
PI	Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI	Daniels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA;
PI	Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI	Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI	Perez-Diaz J, Baquero E, Garcia Del Portillo F, Gomez-Lopez N;
PI	Maduenio E, De Pablos B, Wenland J, Kaerst U, Entian K, Hauf J;
PI	Rose M, Voss H;
XX	
XX	WPI; 2002-010914/01.
XX	
PT	Genomic sequence for <i>Listeria monocytogenes</i> , useful e.g. for treatment
PT	and prevention of <i>Listeria</i> and related bacterial infections, and related
PT	polypeptides.
XX	
PS	Claim 6; SEQ ID NO 2511; 192pp; French.
PS	
XX	
CC	The present invention relates to the genome sequence of <i>Listeria</i>
CC	monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC	it are useful for selecting probes and primers for detecting genes in <i>L.</i>
CC	monocytogenes and related organisms, and for studying genetic
CC	polymorphisms and other genomes. The present sequence is a protein
CC	encoded by the genome sequence of the present invention. Proteins
CC	expressed from the genome sequence are useful for raising specific
CC	antibodies, identification of <i>L. monocytogenes</i> and related organisms, and
CC	for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC	B12. The genome sequence and proteins encoded by it are also useful for
CC	selecting compounds that regulate gene expression and cell replication
CC	and modulate <i>L. monocytogenes</i> -related diseases. In addition, the genome
CC	sequence and proteins encoded by it are useful in pharmaceutical and
CC	vaccines compositions for the treatment or prevention of infections by <i>L.</i>
CC	monocytogenes and related organisms. Note: The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 381 AA:
XX	SO

Query Match	37.7%	Score	703;	DB	5;	Length	381;
Best Local Similarity	38.8%	Pred. No.	2.9e-53;				
Matches	148;	Conservative	93;	Mismatches	134;	Indels	6;
	Gaps	5;					
Qy	2	IQFSINTLFIHALNTTKRAISTKNAJLPIISSIKIEVTSCTGVTLTSGNGOISIEHTPI-V	60				
Db	1	MKFVIEDRLVQAVNEVTRAISSARTTPILTGKILVNDGEGVLTGSDSDISIEAFPLI	60				
Qy	61	SNENAGLLITSPGAILLEASFFINIISLPDISINVKEIEHQHVLTSGKSEITLKGKV	120				
Db	61	ENDEVIVEVESFGGIVLQSKYFGDIVARLPEENVEIEVTSNYQTNISGGQASFTLGLDP	120				
Qy	121	DOYPRLOEVSTENPLILKTKLLKSIIAETAFASLQSRPILITGVHIVLSNHHKDFKAVAT	180				
Db	121	MEYPLPEVTDGKTIKIPINVLKNVIRQTVFVSAIEVRPVLTVGNVMIKENK-LSAVAT	179				
Qy	181	DSHRMSQRLITLD-NTSADLMVLVPLSKSLREFSAVFTDDIETVEVFFSPQILFRSEHIS	239				
Db	180	DSHRLALREIPLTDIDEEYNIVIPKSLSELNKLDDASSIEWTLANNQILPKLKDLL	239				
Qy	240	FYRLLGEGNYPDDRLLMTFETEVVNTQSLRHAMERAFILISNATONGTVKLEITON-H	298				
Db	240	FYSRLSGSPYDTSRLIPTTKSELVINSKAFLOAIDRASLLARENNVNIKMLTLENGQ	299				
Qy	299	ISAHVNSPEGVKNVEDDIDIVSQSGSDITISFNPTYLIESLKAISEVTKTHFLSPVPFT	358				
Db	300	VEVSSNSPEGVNSENVFSQFTGEEIKISFNGKYMMDALRAPGDDIQISFSGTMRPFV	359				
Qy	359	LTPGCD--EESSFIOIITPVRT	377				
Db	360	LRPKDAANPNELIOLITFVRT	380				

RESULT 34
ABU32409

[illegible]

ABU32409 standard; protein; 381 AA.

ABU32409;

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #17936.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Listeria monocytogenes.

WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zye-kind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACA36279.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25: SEQ ID NO 60333; 1766bp: English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences

Query Match 37.7%: Score 703: DB 6: Length 381:

Best Local Similarity 38.8%; Pred. No. 2.9e-53;
Matches 148; Conservative 93; Mismatches 134; Indels 6; Gaps 5;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGOISIENTIP-V 60
Db 1 MKFVIERDRUVQAVNEVTRAIARTIPILTGKIVNDEGVLTGSDSDISTEAFPLI 60
QY 61 SNEGALLITSPGAILLEASFFINISSLIPDISINVKETEQHOVLTSGSEITLKGDOV 120
Db 61 ENDEVIVESEFGIVLQSKYFGDIVRRIPENVEIVTSYQTNISSQASFTLGLDP 120
QY 121 DQYPRIOEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNKHDKFVAT 180
Db 121 MEYFKLPEVTDGKTIPINVLKNIVRQTVFAVAIEVRPLTVGNVILKENK-LSAVAT 179
QY 181 DSHRMSORLITLD-NTSADLMVVLPSKSLREFAVTDIDIEVEVFFSPQILFSEHIS 239
Db 180 DSHRLALRLEIPLETDDEEYNIVIPKSLSELNKLDDASEIEMTLANNQILFKLDLL 239
QY 240 FYTRLLEGNYPDTRLMLTEFEFTEVVFNTQSLRHAMERAFILISNATONCTVKLEITON-H 298
Db 240 FYSRLLEGSPDTRSLIPTDTKSELVINSKAFLOAIDRASLLARENNVILKMTLENGQ 299
QY 299 ISAHVNSPEVGVKNEVDLIVSQSGDLTISFNPTLYIESLKAISKETVKIHLFSPVRPPT 358
Db 300 VEVSNSPEVGVNSVNFQSGFTGEEKISFNGKYMMDALRAFEQGDIDQISFSGTMRPFV 359
QY 359 LTPGD--EESFTQLITPVRT 377
Db 360 LRPKDAANFNEILQLITPVRT 380

RESULT 35
ABU17704
ID ABU17704 standard; protein; 379 AA.
AC ABU17704;
XX
DT 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #3231.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Bacillus anthracis.
OS
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA21574.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 45628; 1766pp; English.
PS
XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 379 AA;

Query Match 37.1%; Score 692; DB 6; Length 379;
Best Local Similarity 38.8%; Pred. No. 2.7e-52;
Matches 148; Conservative 84; Mismatches 141; Indels 8; Gaps 7;
QY 2 IQFSINRTLFHALNTTKRAISTKNAIPILSSIKIEVTSTGVLTGSGOISIENTIPVS 61
Db 1 MRFSIQDYIVRSVQDVVKAVSFRTPIITLTKIKVATEEGVLTGSDADISIESFIPVE 60
QY 62 NENAGLL-ITSPGAILLEASFFINISSLIPDISINVKIEIQHOVV-LTSGKSEITLKGKD 119
Db 61 EDGKEIVEVQSGSIVLQAKYFSEIVKLPKTEVEI-SVENHLMTKITSGKSEFNGLD 119
QY 120 VDQYPRIOEVSTENPLILKTLKLSIAETAFAASLOESRPILTGVHIVLSNKHDKFVAT 179
Db 120 SAEYPLLPQIEEHVFXIPTDLLKHMIRQTVFAVSTSETRPILTVGNWKVYN-SELTCTA 178
QY 180 TDSHRMSORLITLDNTSADLM--VVLPSKSLREFAVTFDIDIEVEVFFSPQILFSEH 237
Db 179 TDSHRLALRKAKEGIVDEFOANVIVPGKSLNELSKILDESEMVDIVITEYQVLFRTKH 238
QY 238 ISFYTRLLEGNYPDTRLMLTEFEFTEVVFNTQSLRHAMERAFILISNATONCTVKLE-ITQ 296
Db 239 LLFFSRLLLEGNYPDTRLIPAESKTDIFVNTKEFLQAI DRASLLARDGRNNVVKLTLEQ 298
QY 297 NHISAHVNSPEVGVKNEVDLIVSQSGDLTISFNPTLYIESLKAISKETVKIHLFSPVRP 356
Db 299 AMLEISNSPEIKGVVEEVOCEKVDGELKISFSAKYMMDALKALDSTEIKISFTGAMRP 358
QY 357 FTLTPGDESESFQILITPVRT 377
Db 359 FLIRTVN-DESIQLILPVRT 378
RESULT 36
ADJ67944
ID ADJ67944 standard; protein; 334 AA.
XX
AC ADJ67944;
XX
DT 20-MAY-2004 (first entry)

```
XX G. stearothermophilus alpha subunit.
DE DNA polymerase; DNA sequencing; DNA amplification.
KW Geobacillus stearothermophilus.
OS US2004038289-A1.
XX 26-FEB-2004.
XX 25-SEP-2003; 2003US-00671403.
XX 08-APR-1997; 97US-0043202P.
PR 08-APR-1998; 98US-00057416.
PR 18-AUG-2000; 2000US-00642218.
PR 21-NOV-2000; 2000US-00716964.
XX (ODON/) O'DONNELL M E.
PA (YUZH/) YUZHAKOV A.
PA (YURI/) YURIEVA O.
PA (JERU/) JERUZALMI D.
PA (BRUC/) BRUCK I.
PA (KURI/) KURIYAN J.
XX O'donnell ME, Yuzhakov A, Yurieva O, Jeruzalmi D, Bruck I;
PI Kuriyan J;
XX WPI; 2004-203218/19.
DR N-PSDB; ADJ67943.
XX Isolated DNA molecule from Bacillus stearothermophilus encoding a delta
PT subunit of a DNA polymerase I mul-type, useful for producing DNA
PT polymerases for use in DNA sequencing and DNA amplification methods.
XX Disclosure; SEQ ID NO 174; 245pp; English.
XX The invention relates to an isolated DNA molecule from Bacillus
CC stearothermophilus encoding a delta subunit of a DNA polymerase I. The
CC subunits are useful for producing DNA polymerases for use in DNA
CC sequencing and DNA amplification methods. The present sequence is used in
CC the exemplification of the present invention.
XX Sequence 334 AA;
Query Match 31.5%; Score 586; DB 8; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.2e-43;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;
QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPPDISINVKIEHQVVL 106
DB 1 NSDISIIESFIPLEKEGKLLVDVKRPGSIVLQARFSEIVKLPQQTVEIETEDNFLTII 60
QY 107 TSGKSEITLKGVDQYPRLOEVSTENPLILKTLKLSIIAETAFAASIQESRPILTVGH 166
DB 61 RSGHSEFRGLNADYPRLOEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSQRLITLDNTSADLM--VVLPSKSLREFSAVFTDD-LETVE 223
DB 121 WKV-ENGELVCTATDSHRLAMRKVKIIESENEVSNVPIPGKSLNELSKIIIDGNGHPVD 179
QY 224 VFPSQILFRSEHISFYTRLLEGNYPOTDRLLMTFETEVVNTQSLRHAMERAFILSN 283
DB 180 IVMATANQVLFKAHLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQADRASLLAR 239
QY 284 ATQNGTVKL-EITQNHISAHVNSPEVGKNEIDLIVSQSGDLTISFNPYILIESLKAIK 342
DB 240 EGRNVNVTTLTPGGMLETSSISPEIGKVTEQLQTESLEGEELNIFSAAKYNMMDALRALD 299
QY 343 SETVKIHFSLSPVRPFTLTGDEESGFILITPVRT 377
DB 300 GTDIQISFTGMRPFLRLP-LHTDSMLQLILPVRT 333
```

```
RESULT 37
ADJ68156
ID ADJ68156 standard; protein; 334 AA.
XX
AC ADJ68156;
XX
DT 20-MAY-2004 (first entry)
XX
DE G. stearothermophilus alpha subunit.
XX
KW DNA polymerase; DNA sequencing; DNA amplification.
XX
OS Geobacillus stearothermophilus.
XX
PN US2004038290-A1.
XX
PD 26-FEB-2004.
XX
PF 25-SEP-2003; 2003US-00671419.
XX
PR 08-APR-1997; 97US-0043202P.
PR 08-APR-1998; 98US-00057416.
PR 18-AUG-2000; 2000US-00642218.
PR 21-NOV-2000; 2000US-00716964.
XX
PA (ODON/) O'DONNELL M E.
PA (YUZH/) YUZHAKOV A.
PA (YURI/) YURIEVA O.
PA (JERU/) JERUZALMI D.
PA (BRUC/) BRUCK I.
PA (KURI/) KURIYAN J.
XX
PI O'donnell ME, Yuzhakov A, Yurieva O, Jeruzalmi D, Bruck I;
PI Kuriyan J;
XX
DR WPI; 2004-203219/19.
DR N-PSDB; ADJ68155.
XX
PT Novel DNA molecule from Thermotoga species encoding delta prime subunit
PT of a DNA polymerase III-type enzyme, useful for producing the enzyme by
PT recombinant techniques.
XX
PS Disclosure; SEQ ID NO 174; 245pp; English.
XX
CC The invention relates to an isolated DNA molecule from Bacillus
CC stearothermophilus encoding a delta subunit of a DNA polymerase I. The
CC subunits are useful for producing DNA polymerases for use in DNA
CC sequencing and DNA amplification methods. The present sequence is used in
CC the exemplification of the present invention.
XX
SQ Sequence 334 AA;
Query Match 31.5%; Score 586; DB 8; Length 334;
Best Local Similarity 37.9%; Pred. No. 5.2e-43;
Matches 127; Conservative 74; Mismatches 126; Indels 8; Gaps 7;
QY 49 NQGIS-IENTIPVSNENAGLL-ITSPGAILLEASFFINIISLPPDISINVKIEHQVVL 106
DB 1 NSDISIIESFIPLEKEGKLLVDVKRPGSIVLQARFSEIVKLPQQTVEIETEDNFLTII 60
QY 107 TSGKSEITLKGVDQYPRLOEVSTENPLILKTLKLSIIAETAFAASIQESRPILTVGH 166
DB 61 RSGHSEFRGLNADYPRLOEENVFQIPADLLKTVIRQTVFAVSTSETRPILTVGN 120
QY 167 IVLSNHKDFKAVATDSHRMSQRLITLDNTSADLM--VVLPSKSLREFSAVFTDD-LETVE 223
DB 121 WKV-ENGELVCTATDSHRLAMRKVKIIESENEVSNVPIPGKSLNELSKIIIDGNGHPVD 179
QY 224 VFPSQILFRSEHISFYTRLLEGNYPOTDRLLMTFETEVVNTQSLRHAMERAFILSN 283
DB 180 IVMATANQVLFKAHLLFFSRLLDGNYPETARLIPTESKTTMIVNAKEFLQADRASLLAR 239
```

Db	61	RS	GHSEFRLNGLNADEYPELPOIESENVFOIPADLLKTVIRQTVFVSTSETRILTVGN	120
Qy	167	IV	LSNHKDFKAVATDSHRMSQRLITLNDTSADLM--VVLPSKSURFSAVFTDD-IETVE	223
Db	121	WKV-EH	GEVLVCTATDSHRLAMEKVKIIESENEVSNVVPKGLNELSKILLDDGNHPVD	179
Qy	224	VFFSPQ	ILFRSHISFYTRLLEGVNPDTDRLLMTPEFETEVFVNTQSLRHAMERAPLISN	283
Db	180	IVMTANQV	LFKAHLLFFSRLLDGNYPETALIPTESKTITMIVNAKEFLQAIDRASLLAR	132
Qy	284	ATQNGTVKL-EITQNH	ISAHVNSPEVGKNEDLDIVSQSGDLTISFNPTYLIESLKAIK	342
Db	240	EGRNNVVKTLL	PGGMLEISSIPISPEIGKVTEQLOQTESLEGEELNIFSAKYMDALRALD	199
Qy	343	SETVKIHELSPVR	PFTLPGDEESSFIOLITPVET	377
Db	300	GTDIQISFTGMR	PFLLRP-LHTDSMLQLILPVET	333
RESULT	39			
ID	ADJ79453			
XX	ADJ79453 standard; protein; 334 AA.			
AC	ADJ79453;			
XX	20-MAY-2004 (first entry)			
DT	XX			
DE	XX			
KW	XX	G. stearothermophilus alpha subunit.		
XX	XX	DNA polymerase; DNA sequencing; DNA amplification.		
OS	XX	Geobacillus stearothermophilus.		
XX	XX	US2004043414-A1.		
PN	XX	04-MAR-2004.		
PD	XX	25-SEP-2003; 2003US-00670844.		
PF	XX	08-APR-1997; 97US-0043202P.		
PR	XX	08-APR-1998; 98US-00057416.		
PR	XX	18-AUG-2000; 2000US-00642218.		
PR	XX	21-NOV-2000; 2000US-00716964.		
XX	XX	(ODON/) O'DONNELL M E.		
PA	XX	(YUZH/) YUZHAKOV A.		
PA	XX	(YURI/) YURIEVA O.		
PA	XX	(JERU/) JERUZALMI D.		
PA	XX	(BRUC/) BRUCK I.		
PA	XX	(KURI/) KURIYAN J.		
XX	XX	O'donnell ME, Yuzhakov A, Yurieva O, Jeruzalmi D, Bruck I;		
PI	XX	Kuriyan J;		
XX	XX	WPI; 2004-225698/21.		
DR	XX	N-PSDB; ADJ79452.		
XX	XX			
PT	PT	Novel isolated DNA molecule from Bacillus stearothermophilus, encoding		
PT	PT	tau subunit of DNA polymerase III-type enzyme, useful in amplification		
PT	PT	and sequencing reactions.		
PS	PS	Disclosure; SEQ ID NO 174; 245pp; English.		
CC	CC	The invention relates to an isolated DNA molecule from Bacillus		
CC	CC	stearothermophilus encoding a delta subunit of a DNA polymerase I. The		
CC	CC	subunits are useful for producing DNA polymerases for use in DNA		
CC	CC	sequencing and DNA amplification methods. The present sequence is used in		
CC	CC	the exemplification of the present invention.		
XX	XX	Sequence 334 AA;		
SQ				

